



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification: C12N 15/86, C07K 14/16, C12N 5/10, C12N 7/04, C12N 15/49	A1	(11) International Publication Number: WO 00/15819 (43) International Publication Date: 23 March 2000 (23.03.2000)
(21) International Application Number: PCT/US99/20675 (22) International Filing Date: 10 September 1999 (10.09.1999)		Published
(30) Priority Data: 60/100,022 11 September 1998 (11.09.1998) US 60/100,063 12 September 1998 (12.09.1998) US		
(60) Parent Application or Grant THE CHILDREN'S MEDICAL CENTER CORPORATION []; O. GRAY, John, T. []; O. MULLIGAN, Richard, C. []; O. BROOK, David, E. ; O.		
(54) Title: PACKAGING CELL LINES FOR HIV-DERIVED RETROVIRAL VECTOR PARTICLES (54) Titre: LIGNEES DE CELLULES D'ENCAPSIDATION POUR PARTICULES DE VECTEUR RETROVIRAL DERIVE DU VIH		
(57) Abstract Novel packaging cell lines useful for generating viral accessory protein independent HIV-derived retroviral vector particles, methods of constructing such packaging cell lines and methods of using the viral accessory protein independent HIV-derived retroviral vector particles are disclosed.		
(57) Abrégé L'invention concerne de nouvelles lignées de cellules d'encapsidation utiles pour produire des particules de vecteur rétroviral dérivé du VIH indépendantes de protéines accessoires virales, des procédés de mise au point de ces lignées de cellules d'encapsidation et des procédés d'utilisation des particules de vecteur rétroviral dérivé du VIH indépendantes de protéines accessoires virales.		

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 15/86, 5/10, 7/04, 15/49, C07K 14/16	A1	(11) International Publication Number: WO 00/15819 (43) International Publication Date: 23 March 2000 (23.03.00)		
(21) International Application Number: PCT/US99/20675		(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).		
(22) International Filing Date: 10 September 1999 (10.09.99)				
(30) Priority Data: 60/100,022 11 September 1998 (11.09.98) US 60/100,063 12 September 1998 (12.09.98) US		Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>		
(71) Applicant: THE CHILDREN'S MEDICAL CENTER CORPORATION (US/US); 300 Longwood Avenue, Boston, MA 02115 (US).				
(72) Inventors: GRAY, John, T.; 48 Spring Road, West Roxbury, MA 02132 (US). MULLIGAN, Richard, C.; 2 Sandy Pond Road, Lincoln, MA 01773 (US).				
(74) Agents: BROOK, David, E. et al.; Hamilton, Brook, Smith & Reynolds, P.C., Two Militia Drive, Lexington, MA 02421 (US).				
(54) Title: PACKAGING CELL LINES FOR HIV-DERIVED RETROVIRAL VECTOR PARTICLES				
<p>The diagram illustrates the structure of HIV provirus and three packaging cell lines. HIV provirus is shown with its genome components: LTR, GAG, PRO, POL, VIF, R, TAT, REV, ENV, RRE, and LTR. The pCMV AR8.2 vector contains the HIV GAG, PRO, POL, VIF, R, TAT, REV, ENV, RRE, and LTR regions, along with a CMV promoter (SD) driving the expression of the LacZ gene. The pHR'-CMV-LacZ vector is similar but lacks the REV gene. The pMD.G vector contains the CMV promoter (SD), the VSVG envelope protein gene, and the polyA signal sequence.</p>				
<p>(Naldini et al, <i>Science</i> 272:263, 1996)</p>				
(57) Abstract				
<p>Novel packaging cell lines useful for generating viral accessory protein independent HIV-derived retroviral vector particles, methods of constructing such packaging cell lines and methods of using the viral accessory protein independent HIV-derived retroviral vector particles are disclosed.</p>				

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Angola	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Marshall Islands	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NB	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KB	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		

Description

5

10

15

20

25

30

35

40

45

50

55

5

-1-

10

PACKAGING CELL LINES FOR HIV-DERIVED RETROVIRAL VECTOR PARTICLES

BACKGROUND OF THE INVENTION

15 Retroviral vectors based on lentiviruses, such as human immunodeficiency viruses (HIV), can infect nondividing cells, and integration of proviral DNA occurs
20 5 without the need for cell division. These properties make lentiviruses attractive for gene transfer into nondividing cells, such as hepatocytes, myofibers, hematopoietic stem cells, and neurons.

25 However, the use of lentivirus vectors, particularly HIV vectors, particularly for gene therapy, is hampered by concern over their safety. Thus, a need for the
30 10 development of lentivirus vectors, particularly HIV vectors, with improved safety, particularly for gene therapy, exists.

SUMMARY OF THE INVENTION

35 The present invention relates to novel packaging cell lines useful for generating viral accessory protein independent lentivirus-derived, particularly HIV-derived,
40 15 retroviral vector particles, to construction of such cell lines and to methods of using the accessory protein independent lentivirus-derived retroviral vector particles to introduce DNA of interest into cells (e.g., eukaryotic cells such as animal (particularly mammalian), plant or yeast cells or prokaryotic cells such as bacterial cells). In a preferred embodiment, the packaging cell lines of the present invention are stable
45 20 packaging cell lines.

50 In one embodiment of the invention, packaging cell lines for producing a viral accessory protein independent lentivirus-derived retroviral vector particles comprise (a) a cell (e.g., mammalian cell); and (b) a retroviral nucleotide sequence in the cell which comprises a coding sequence for lentivirus *gagpol*, wherein said coding sequence has

55

5

-2-

10

been codon optimized by mutagenesis to improve expression of the lentivirus gagpol proteins.

15

In second embodiment of the invention, packaging cell lines for producing a viral accessory protein independent lentivirus-derived retroviral vector particles comprise (a) a cell (e.g., mammalian cell); (b) a first retroviral nucleotide sequence in the cell which comprises a coding sequence for lentivirus *gagpol*, wherein said coding sequence has been codon optimized by mutagenesis to improve expression of the lentivirus gagpol proteins; and (c) a second retroviral nucleotide sequence in the cell which comprises the coding sequence for a heterologous envelope protein.

20

10 In a third embodiment of the invention, packaging cell lines for producing a viral accessory protein independent lentivirus-derived retroviral vector particles comprise (a) a cell (e.g., mammalian cell); (b) a first retroviral nucleotide sequence in the cell which comprises a coding sequence for lentivirus *gagpol*, wherein said coding sequence has been codon optimized by mutagenesis to improve expression of the lentivirus gagpol proteins; (c) a second retroviral nucleotide sequence in the cell which comprises the coding sequence for a heterologous envelope protein; and (d) a third retroviral nucleotide sequence which comprises a DNA sequence of interest and lentivirus cis-acting sequences required for packaging, reverse transcription and integration.

25

30

35 In a fourth embodiment of the invention, packaging cell lines for producing a viral accessory protein independent lentivirus-derived retroviral vector particles comprise (a) a cell (e.g., mammalian cell); (b) a retroviral nucleotide sequence in the cell which comprises a coding sequence for lentivirus *gagpol*, wherein said coding sequence has been codon optimized by mutagenesis to improve expression of the lentivirus gagpol proteins; and (c) a retroviral nucleotide sequence which comprises a 40 DNA sequence of interest and lentivirus cis-acting sequences required for packaging, reverse transcription and integration.

45

45 In a fifth embodiment of the invention, packaging cell lines for producing a viral accessory protein independent HIV-derived retroviral vector particles comprise (a) a cell

50

55

5

-3-

10 (e.g., mammalian cell); and (b) a retroviral nucleotide sequence in the cell which comprises a coding sequence for HIV *gagpol*, wherein said coding sequence has been codon optimized by mutagenesis to improve expression of the HIV gagpol proteins.

15 In sixth embodiment of the invention, packaging cell lines for producing a viral accessory protein independent HIV-derived retroviral vector particles comprise (a) a cell (e.g., mammalian cell); (b) a first retroviral nucleotide sequence in the cell which comprises a coding sequence for HIV *gagpol*, wherein said coding sequence has been codon optimized by mutagenesis to improve expression of the HIV gagpol proteins; and (c) a second retroviral nucleotide sequence in the cell which comprises the coding 20 sequence for a heterologous envelope protein.

25 In a seventh embodiment of the invention, packaging cell lines for producing a viral accessory protein independent HIV-derived retroviral vector particles comprise (a) a cell (e.g., mammalian cell); (b) a first retroviral nucleotide sequence in the cell which comprises a coding sequence for HIV *gagpol*, wherein said coding sequence has been 30 codon optimized by mutagenesis to improve expression of the HIV gagpol proteins; (c) a second retroviral nucleotide sequence in the cell which comprises the coding sequence for a heterologous envelope protein; and (d) a third retroviral nucleotide sequence which comprises a DNA sequence of interest and HIV cis-acting sequences required for 35 packaging, reverse transcription and integration.

40 20 In a eighth embodiment of the invention, packaging cell lines for producing a viral accessory protein independent HIV-derived retroviral vector particles comprise (a) a cell (e.g., mammalian cell); (b) a retroviral nucleotide sequence in the cell which comprises a coding sequence for HIV *gagpol*, wherein said coding sequence has been codon optimized by mutagenesis to improve expression of the HIV gagpol proteins; and 45 (c) a retroviral nucleotide sequence which comprises a DNA sequence of interest and HIV cis-acting sequences required for packaging, reverse transcription and integration.

50 Alternatively, each of the packaging cell lines described herein can be produced using (1) a retroviral nucleotide sequence which comprises a codon optimized gag

55

5

-4-

10 coding sequence and (2) a retroviral nucleotide sequence which comprises a codon optimized pol coding sequence, in place of the retroviral nucleotide sequence which comprises a codon optimized gagpol coding sequence.

15 In a particular embodiment, the heterologous envelope protein is the G glycoprotein of vesicular stomatitis virus (VSV G). In another embodiment, the heterologous envelope protein is the amphotropic envelope of the Moloney leukemia virus (MLV).

20 Cell lines for producing a viral accessory protein independent lentivirus-derived retroviral vector particles are produced by transfecting host cells (e.g., mammalian host 25 cells) with a plasmid comprising a DNA sequence which encodes lentivirus *gagpol* proteins, wherein said DNA sequence has been codon optimized by mutagenesis to improve expression of the lentivirus gagpol proteins. Depending upon the particular cell line being produced, the host cells are also co-transfected with a plasmid comprising a DNA sequence which encodes a heterologous envelope protein, or a 30 plasmid comprising a DNA sequence of interest and lentivirus *cis*-acting sequences required for packaging, reverse transcription and integration, or both of these plasmids. Alternatively, host cells are transfected with a plasmid comprising a codon optimized DNA sequence encoding a lentivirus gag protein and a plasmid comprising a codon 35 optimized DNA sequence encoding a lentivirus pol protein, in place of the plasmid 20 comprising a codon optimized DNA sequence encoding both lentivirus gagpol proteins.

40 Cell lines for producing a viral accessory protein independent HIV-derived retroviral vector particles are produced by co-transfecting host cells (e.g., mammalian host cells) with a plasmid comprising a DNA sequence which encodes HIV *gagpol* proteins, wherein said DNA sequence has been codon optimized by mutagenesis to 45 improve expression of the HIV gagpol proteins. Depending upon the particular cell line being produced, the host cells are also co-transfected with a plasmid comprising a DNA sequence which encodes a heterologous envelope protein, or a plasmid comprising a DNA sequence of interest and HIV *cis*-acting sequences required for packaging, reverse

50

55

5

-5-

transcription and integration, or both of these plasmids. Alternatively, host cells are
10 transfected with a plasmid comprising a codon optimized DNA sequence encoding a
HIV gag protein and a plasmid comprising a codon optimized DNA sequence encoding
a HIV pol protein, in place of the plasmid comprising a codon optimized DNA sequence
15 encoding both HIV gagpol proteins.

The present invention also relates to methods of producing viral accessory
protein independent lentivirus-derived retroviral vector particles, comprising co-
transfected host cells (e.g., mammalian host cells) with (a) a first plasmid comprising a
20 DNA sequence which encodes lentivirus gagpol proteins, wherein said DNA sequence
has been codon optimized by mutagenesis to improve expression of the lentivirus gagpol
proteins; (b) a second plasmid comprising a DNA sequence which encodes a
25 heterologous envelope protein; and (c) a third plasmid comprising a DNA sequence of
interest and lentivirus cis-acting sequences required for packaging, reverse transcription
and integration. Alternatively, host cells are transfected with a plasmid comprising a
30 codon optimized DNA sequence encoding a lentivirus gag protein and a plasmid
comprising a codon optimized DNA sequence encoding a lentivirus pol protein, in place
of the first plasmid comprising a codon optimized DNA sequence encoding both
lentivirus gagpol proteins.

In a particular embodiment, the invention relates to methods of producing viral
35 accessory protein independent HIV-derived retroviral vector particles, comprising co-
transfected host cells (e.g., mammalian host cells) with (a) a first plasmid comprising a
DNA sequence which encodes HIV gagpol proteins, wherein said DNA sequence has
40 been codon optimized by mutagenesis to improve expression of the HIV gagpol
proteins; (b) a second plasmid comprising a DNA sequence which encodes a
45 heterologous envelope protein; and (c) a third plasmid comprising a DNA sequence of
interest and HIV cis-acting sequences required for packaging, reverse transcription and
integration. Alternatively, host cells are transfected with a plasmid comprising a codon
optimized DNA sequence encoding a HIV gag protein and a plasmid comprising a

50

55

5

-6-

10 codon optimized DNA sequence encoding a HIV pol protein, in place of the first plasmid comprising a codon optimized DNA sequence encoding both HIV gagpol

proteins.

15 The present invention also relates to viral accessory protein-independent 5 retroviral particles produced by or obtainable by (obtained by) the methods described herein.

20 The present invention further relates to isolated DNA encoding a codon 10 optimized lentivirus *gagpol*, isolated DNA encoding the *gag* coding region of a codon optimized lentivirus *gagpol*, and isolated DNA encoding the *pol* coding region of a 25 10 codon optimized lentivirus *gagpol*. In a particular embodiment, the present invention relates to isolated DNA encoding a codon optimized HIV *gagpol*, isolated DNA encoding the *gag* coding region of a codon optimized HIV *gagpol*, and isolated DNA encoding the *pol* coding region of a codon optimized HIV *gagpol*.

30 The packaging cell lines and viral particles of the present invention can be used 15 for gene therapy or gene replacement with improved safety. The packaging cell lines and viral particles of the present invention can also be used in development and production of vaccines, and in production of biochemical reagents. Gene therapy vectors produced with the cell lines of the present invention are expected to be valuable 35 medical therapeutics.

20 BRIEF DESCRIPTION OF THE DRAWINGS

40 Figure 1 is a schematic diagram of an expression cassette containing the codon 45 optimized *gagpol* genes. The DNA was constructed in multiple segments, which are indicated at the top as 1/3, 2/3, 3/3 (A, B, C and D) and HIN. Restriction sites used to assemble the cloned segments are indicated above the kilobasepair (Kb) ruler. Below 50 25 the ruler are multiple features showing the location of the human cytomegalovirus (CMV) promoter, human betaglobin sequences (Bglobin), mRNA sequences (thinner line represents intronic sequence), the *gag* and *pol* open reading frames, the individual

55

5

-7-

10 proteolytic fragment coding sequences (p17_MA, p24_CA, p7, p6, PR, p51_RT, RNaseH and integrase (IN)) and each synthetic oligonucleotide used in the assembly

process (multiple adjacent open arrows).

15 Figure 2 is a table which depicts codon usage frequencies in genes which are highly expressed and in the codon optimized gagpol open reading frame of the HIV packaging construct described herein.

20 Figure 3 is a schematic representation of the HIV provirus and a three-plasmid expression system used for generating a pseudotyped HIV-based vector by transient transfection as described in Naldini *et al.*, *Science*, 272:263-267 (1996).

25 10 Figure 4 is a list of some characteristics relating to the HIV Rev protein.

Figure 5 is a list of some points relating to codon optimization of HIV *gagpol*.

30 25 Figure 6 is a partial DNA sequence of HIV *gag* (SEQ ID NO: 1), showing inactivation of inhibitory sequences as described in Schwartz, S. *et al.*, *J. Virol.*, 66(12):7176-7182 (1992).

35 15 Figure 7 a plot of the %(G+C) content of wildtype HIV *gagpol* sequences and theoretically codon optimized HIV *gagpol* sequences. The percent of bases, either G or C, was calculated for a 30 nucleotide moving window for the entire length of the *gagpol* gene, and the value plotted versus nucleotide position. Diamonds = HIV *gagpol* sequences; squares = full optimal back-translation for *gag* open reading frame; 20 triangles = full optimal back-translation for *pol* open reading frame; CO = codon optimized.

40 40 Figures 8A-8E depict the alignment of the nucleotide sequences and predicted amino acid sequences for the *gag* coding region of a wildtype HIV *gagpol* and a codon optimized HIV *gagpol*. "NL4-3 genbank.SEQ" indicates the nucleotide sequence (SEQ 45 25 ID NO:2) and predicted amino acid sequence (SEQ ID NO:3) for the *gag* coding region of a wildtype HIV *gagpol*. "pHDMHgpM2.seq" indicates the nucleotide sequence (SEQ ID NO:4) and predicted amino acid sequence (SEQ ID NO:5) for the *gag* coding region

50

55

5

-8-

10

of a codon optimized HIV *gagpol*. The "NL4-3 genbank.SEQ" sequences are publicly available at the NIH GenBank sequence repository (Accesssion No. M19921).

15

Figures 9A-9L depict the alignment of the nucleotide sequences and predicted amino acid sequences for the *pol* coding region of a wildtype HIV *gagpol* and a codon optimized HIV *gagpol*. "NL4-3 genbank.SEQ" indicates a nucleotide sequence (SEQ ID NO:6) and a predicted amino acid sequence (SEQ ID NO:7) for the *pol* coding region of a wildtype HIV *gagpol* available in the NIH GenBank sequence repository (Accesssion No. M19921). The nucleotide and amino acid sequences for the *pol* coding region available in the GenBank sequence repository contain two sequence errors, which are indicated in Figures 9A-9L with shading. "pNL4-3.seq" indicates the correct nucleotide sequence (SEQ ID NO:8) and predicted amino acid sequence (SEQ ID NO:9) for the *pol* coding region of a wildtype HIV *gagpol*. "pHDMHgpm2.seq" indicates the nucleotide sequence (SEQ ID NO:10) and predicted amino acid sequence (SEQ ID NO:11) for the *pol* coding region of a codon optimized HIV *gagpol*.

20

Figures 10A-10D depict the DNA sequence (SEQ ID NO:12) for pHDMHgpm2. The CMV enhancer/promoter is at nucleotides 97 to 679, human betaglobin sequences (Bglobin) are at nucleotides 761 to 864, 865 to 1303 and 5710 to 6469 (end of Bglobin is at nucleotides 6445 to 6469), mRNA sequences are at nucleotides 680 to 778 and 1255 to 5921, SV40 origin of replication is at nucleotides 8796 to 8908, beta-lactamase (bla) coding region is at nucleotides 6709 to 7569, intron sequences are at nucleotides 779 to 1254, the codon optimized *gag* coding region is at nucleotides 1318 to 2820, the codon optimized *pol* coding region is at nucleotides 2619 to 5624 and the poly A site is at nucleotides 5897 to 5921.

25

Figure 11 is a circular map of plasmid pHDMHgpm2.

30

25 DETAILED DESCRIPTION OF THE INVENTION

35

The present invention relates to novel packaging cell lines useful for generating viral accessory protein independent lentivirus-derived, particularly HIV-derived,

40

45

50

55

5

-9-

10

retroviral vector particles, to construction of such cell lines and to methods of using the
accessory protein independent lentivirus-derived retroviral vector particles to introduce
DNA of interest into cells (e.g., eukaryotic cells such as animal (particularly
mammalian), plant or yeast cells or prokaryotic cells such as bacterial cells). In a
5 particular embodiment, the packaging cell lines of the present invention are stable
15 packaging cell lines.

15

The cell lines are engineered to express the lentivirus proteins necessary for
virus particle formation (*gagpol* proteins), without containing DNA sequences from
20 lentivirus accessory proteins (*tat*, *vif*, *vpr*, *vpu*, *nef* and *rev* proteins and Rev response
10 element (RRE)). Additionally, no viral sequences (such as cis-acting elements termed
constitutive transport elements (CTEs)) will be expressed as RNA of any kind. DNA
sequences for lentivirus *gagpol* are codon optimized by extensively mutagenizing the
25 sequences to improve expression and to reduce the risk of recombination between
transfer vector sequences and *gagpol* messenger RNA. This greatly improves the safety
15 of virus preparations generated from these cell lines. In a particular embodiment, the
30 DNA sequences for lentivirus *gagpol* are not codon optimized in the overlap region
between the *gag* and *pol* sequences and in cis-acting signals necessary for translation of
pol.

35

Examples of lentiviruses include human immunodeficiency viruses (e.g., HIV-1,
20 HIV-2, HIV-3), bovine lentiviruses (e.g., bovine immunodeficiency viruses, bovine
immunodeficiency-like viruses, Jembrana disease viruses), equine lentiviruses (e.g.,
40 equine infectious anemia viruses), feline lentiviruses (e.g., feline immunodeficiency
viruses, panther lentiviruses, puma lentiviruses), ovine/caprine lentiviruses (e.g.,
Brazilian caprine lentiviruses, caprine arthritis-encephalitis viruses, Maedi-Visna
45 viruses, Maedi-Visna-like viruses, Maedi-Visna-related viruses, ovine lentiviruses,
Visna lentiviruses), Simian AIDS retroviruses (e.g., human T-cell lymphotropic virus
type 4), simian immunodeficiency viruses, simian-human immunodeficiency viruses,
human lymphotrophic viruses (e.g., type III), simian T-cell lymphotrophic viruses.

50

55

5

10

In another embodiment, cell lines are engineered to express the HIV proteins necessary for virus particle formation (*gagpol* proteins), without containing DNA sequences from HIV accessory proteins (*lat*, *vif*, *vpr*, *vpu*, *nef* and *rev* proteins and Rev response element (RRE)). Additionally, no viral sequences (such as cis-acting elements 5 termed constitutive transport elements (CTEs)) will be expressed as RNA of any kind. DNA sequences for a HIV *gagpol* are codon optimized by mutagenesis to improve expression and to reduce the risk of recombination between transfer vector sequences and *gagpol* messenger RNA. In a particular embodiment, the DNA sequences for HIV *gagpol* are not codon optimized in the overlap region between the *gag* and *pol* 10 sequences and in cis-acting signals necessary for translation of *pol*.

Alternatively, each of the packaging cell lines described herein can be produced using (1) a nucleotide sequence which comprises a codon optimized gag coding sequence and (2) a nucleotide sequence which comprises a codon optimized pol coding sequence, in place of the nucleotide sequence which comprises a codon optimized gagpol coding sequence. In this embodiment, the gag and pol coding sequences can be completely codon optimized.

Benefits of the present invention include the removal of potentially harmful lentivirus accessory proteins and other viral sequences, and the reduction of the risk of recombination to produce replication competent virus.

20 Packaging cell lines for producing a viral accessory protein independent
lentivirus-derived retroviral vector particles comprise a mammalian cell and a retroviral
nucleotide sequence comprising a coding sequence for a lentivirus *gagpol* which has
been codon optimized. In a particular embodiment the packaging cell lines further
comprise a retroviral nucleotide sequence comprising a coding sequence for a
25 heterologous envelope protein. In a second embodiment, the packaging cell lines
further comprise a retroviral nucleotide sequence comprising a coding sequence for a
heterologous envelope protein and a retroviral nucleotide sequence which comprises a
DNA sequence of interest and HIV cis-acting sequences required for packaging, reverse

50

55

5

-11-

transcription and integration. In third embodiment, the packaging cell lines further
comprise a retroviral nucleotide sequence which comprises a DNA sequence of interest
and HIV cis-acting sequences required for packaging, reverse transcription and
integration. Alternatively, the packaging cell lines of the present invention comprise a
retroviral nucleotide sequence which comprises a codon optimized gag coding sequence
and (2) a retroviral nucleotide sequence which comprises a codon optimized pol coding
sequence, in place of the retroviral nucleotide sequence which comprises a codon
optimized gagpol coding sequence.

20

The coding sequence(s) for lentivirus *gagpol* which has (have) been codon

10 optimized results in improved expression of the lentivirus gagpol proteins and reduces
the risk of recombination between the transfer vector and gagpol messenger RNA.
Codon optimization of the coding sequence(s) for lentivirus *gagpol* was obtained by
mutagenizing for each particular amino acid residue, specific nucleic acid bases in a
codon for the particular amino acid residue to a nucleic acid base which is present in a
codon which occurs at a high frequency in genes which are highly expressed for the
same amino acid residue. In a particular embodiment, the resulting optimized codon
also does not cause introduction of mRNA splicing signals into the codon optimized
sequence. Thus, in a particular embodiment, codon optimization of the coding
sequence(s) for lentivirus *gagpol* is obtained by mutagenizing for each particular amino
acid residue, specific nucleic acid bases in a codon for the particular amino acid residue
to a nucleic acid base that is present in a codon which (1) occurs at a high frequency in
genes which are highly expressed for the same amino acid residue and (2) does not
cause introduction of mRNA splicing signals into the codon optimized sequence.
Codon optimization typically results in the removal of nucleic acid base A-rich
instability elements.

45

In a particular embodiment, the coding sequence for a HIV *gagpol* (pNL4-3;
available through the AIDS repository, NIH; Adachi *et al.*, *J. Virol.*, 59:284-291 (1986))
has been codon optimized to improve translational efficiency of the HIV gagpol

50

55

5

-12-

proteins and reduce the risk of recombination between the transfer vector and HIV
10 gagpol messenger RNA. Two hundred thirty-seven base pairs (237 bp) consisting of the
gag pol overlap and cis-acting signals necessary for translation of pol (nucleotides 2583
to 2819 of SEQ ID NO: 12) were not optimized. The HIV *gagpol* sequence obtained—
15 5 using the codon optimization process does not differ at the amino acid level from the
wildtype HIV *gagpol* sequence, but differs at the nucleotide level from the HIV *gagpol*
sequence. A codon optimized HIV *gag* sequence is shown in Figures 8A-8E
(pHDMHgpm2.seq) (SEQ ID NO:4). A codon optimized HIV *pol* sequence is shown in
20 Figures 9A-9L (pHDMHgpm2.scq) (SEQ ID NO:10).

10 10 A plasmid comprising DNA sequences which encode codon optimized lentivirus
gagpol proteins is also referred to herein as a packaging construct. This plasmid
includes a promoter which drives the expression of the gagpol proteins, such as the
25 human cytomegalovirus (hCMV) immediate early promoter. This plasmid is defective
for the production of the viral envelope and accessory proteins tat, vif, vpr, vpu, nef and
15 rev and the Rev response element (RRE). The packaging construct also does not
contain viral sequences which are transcribed into mRNA, such as constitutive transport
30 elements (CTEs).

A packaging construct comprising a codon optimized HIV *gagpol* is depicted in
35 Figure 1 and in Figure 11. Figures 10A-10D depict the DNA sequence (SEQ ID
NO:12) for the packaging construct pHDMHgpm2. This packaging construct
20 (pHDMHgpm2) was constructed as follows: Plasmid pMDA.HIVgp mam was
generated by chemical synthesis and PCR assembly (which is described in, for example,
40 Stemmer *et al.*, *Gene*, 164:49-53 (1995)) of 215 different oligonucleotides. The DNA
sequence for pMDA.HIVgp mam is the same as the DNA sequence for
25 pMDA.HIVgp jtg except for 4.3 kb which was codon optimized using the DNASTar
program (LaserGene, Madison, WI). Two hundred thirty-seven base pairs (237 bp)
45 consisting of the gag pol overlap and cis-acting signals necessary for translation of pol
(nucleotides 2583 to 2819 of SEQ ID NO: 12) were not optimized due to dual reading.

50

55

5

-13-

frame constraints. A Nsil site 5' of IN was preserved to aid fusion with wildtype sequences. Several single or double base pair silent mutations were introduced either to prevent potential splice donors and acceptors, or by the synthesis process.

10 pMDA.HIV gp jtg was derived from HIV-1 strain NL4-3. The protease mutation that is present in the NL4-3 NIH GenBank sequence was then repaired (Figure 9B), changing the nucleotide present at position 2948 of SEQ ID NO:12 from a "G" to a "C", thereby producing the codon present at nucleotide positions 2948 to 2950 of SEQ ID NO:12 which encodes an arginine instead of the glycine present in the NL4-3 GenBank amino acid sequence. The resulting plasmid was named pMDHgpmam. The EcoRI-HindIII 15 fragment of pMDHgpmam was inserted into pHDM2b, a high copy version of the pMD vector (Ory, D. et al., *Proc. Natl. Acad. Sci. USA*, 93(21):11400-11406 (1996)), to produce plasmid pHDMHgpm. The sequencing mutation that is present in the RNase domain of the NL4-3 NIH GenBank sequence was repaired (Figure 9H), changing the nucleotide present at position 4724 of SEQ ID NO:12 from "GGG" to 20 "AAG", thereby producing a codon encoding a lysine instead of the glycine present in the NL4-3 GenBank amino acid sequence. The resulting plasmid was named pHDMHgpm2. Codon usage frequencies in the codon optimized gagpol open reading frame of the packaging construct pHDMHgpm2 are shown in Figure 2.

25 As used herein, a heterologous envelope protein permits pseudotyping of particles generated by the packaging construct and includes the G glycoprotein of vesicular stomatitis virus (VSV G) and the amphotropic envelope of the Moloney leukemia virus (MLV). A plasmid comprising a DNA sequence which encodes a 30 heterologous envelope protein is also referred to herein as an envelope coding plasmid.

35 The terms "mammal" and "mammalian", as used herein, refer to any vertebrate animal, including monotremes, marsupials and placental, that suckle their young and 40 either give birth to living young (eutherian or placental mammals) or are egg-laying (metatharian or nonplacental mammals). Examples of mammalian species include 45

50

55

5

-14-

10

humans and other primates (e.g., monkeys, chimpanzees), rodents (e.g., rats, mice, guinea pigs) and ruminants (e.g., cows, pigs, horses).

15

Examples of mammalian cells include human (such as HeLa cells, 293T cells, NIH 3T3 cells), bovine, ovine, porcine, murine (such as embryonic stem cells), rabbit and monkey (such as COS1 cells) cells. The cell may be a non-dividing cell (including hepatocytes, myofibers, hematopoietic stem cells, neurons) or a dividing cell. The cell may be an embryonic cell, bone marrow stem cell or other progenitor cell. Where the cell is a somatic cell, the cell can be, for example, an epithelial cell, fibroblast, smooth muscle cell, blood cell (including a hematopoietic cell, red blood cell, T-cell, B-cell, etc.), tumor cell, cardiac muscle cell, macrophage, dendritic cell, neuronal cell (e.g., a glial cell or astrocyte), or pathogen-infected cell (e.g., those infected by bacteria, viruses, virusoids, parasites, or prions).

20

Typically, cells isolated from a specific tissue (such as epithelium, fibroblast or hematopoietic cells) are categorized as a "cell-type." The cells can be obtained commercially or from a depository or obtained directly from an animal, such as by biopsy. Alternatively, the cell need not be isolated at all from the animal where, for example, it is desirable to deliver the virus to the animal in gene therapy.

25

To produce the cell lines of the present invention for producing a viral accessory protein independent lentivirus-derived retroviral vector particles, mammalian host cells are co-transfected with (a) a first plasmid comprising DNA sequence which encode lentivirus gagpol proteins, wherein said DNA sequence has been codon optimized by mutagenesis, as described above, to improve expression of the lentivirus gagpol proteins; and (2) a second plasmid comprising a DNA sequence which encodes a heterologous envelope protein, or a retroviral nucleotide sequence which comprises a DNA sequence of interest and lentivirus cis-acting sequences required for packaging, reverse transcription and integration, or both, under conditions appropriate for transfection of the cells.

40

45

50

55

5

-15-

In a particular embodiment, to produce the cell lines of the present invention for
10 producing viral accessory protein independent HIV-derived retroviral vector particles
mammalian host cells were cotransfected with (a) a first plasmid comprising DNA
sequence which encode HIV *gagpol* proteins, wherein said DNA sequence has been
15 codon optimized by mutagenesis, as described above, to improve expression of the HIV
gagpol proteins; and (2) a second plasmid comprising a DNA sequence which encodes a
heterologous envelope protein, or a retroviral nucleotide sequence which comprises a
DNA sequence of interest and HIV cis-acting sequences required for packaging, reverse
20 transcription and integration, or both, under conditions appropriate for transfection of
the cells.

Virus stocks consisting of viral accessory protein independent lentivirus-derived,
25 particularly HIV-derived, retroviral vector particles of the present invention are
produced by maintaining the transfected cells under conditions suitable for virus
production (e.g., in an appropriate growth media and for an appropriate period of time).
30 Such conditions, which are not critical to the invention, are generally known in the art.
See, e.g., *Sambrook et al., Molecular Cloning: A Laboratory Manual*, Second Edition,
Cold Spring Harbor University Press, New York (1989); *Ausubel et al., Current
Protocols in Molecular Biology*, John Wiley & Sons, New York (1998); U.S. Patent
35 No. 5,449,614; and U.S. Patent No. 5,460,959; the teachings of which are incorporated
herein by reference.

To generate viral accessory protein independent lentivirus-derived retroviral
40 vector particles, mammalian host cells can be co-transfected with (a) a first plasmid
comprising DNA sequence which encode lentivirus *gagpol* proteins, wherein said DNA
sequence has been codon optimized by mutagenesis, as described above, to improve
45 expression of the lentivirus *gagpol* proteins; (b) a second plasmid comprising a DNA
sequence which encodes a heterologous envelope protein; and (c) a third plasmid
comprising a DNA sequence of interest and lentivirus cis-acting sequences required for
packaging, reverse transcription and integration. Alternatively, mammalian cells are

50

55

5

-16-

transfected with a plasmid comprising a codon optimized DNA sequence encoding a
10 lentivirus gag protein and a plasmid comprising a codon optimized DNA sequence
encoding a lentivirus pol protein, in place of the first plasmid comprising a codon
optimized DNA sequence encoding both lentivirus gagpol proteins. Alternatively,

15 5 mammalian host cells are transfected with a plasmid comprising a codon optimized
DNA sequence encoding a lentivirus gag protein and a plasmid comprising a codon
optimized DNA sequence encoding a lentivirus pol protein, in place of the first plasmid
comprising a codon optimized DNA sequence encoding both lentivirus gagpol proteins.

20 In a particular embodiment, the invention relates to methods of producing viral
accessory protein independent HIV-derived retroviral vector particles, comprising co-
25 transfecting mammalian host cells with (a) a first plasmid comprising DNA sequence
which encode HIV gagpol proteins, wherein said DNA sequence has been codon
optimized by mutagenesis, as described above, to improve expression of the HIV gagpol
proteins; (b) a second plasmid containing a DNA sequence which encodes a
30 15 heterologous envelope protein; and (c) a third plasmid comprising a DNA sequence of
interest and HIV cis-acting sequences required for packaging, reverse transcription and
integration. Alternatively, mammalian host cells are transfected with a plasmid
comprising a codon optimized DNA sequence encoding a HIV gag protein and a
35 20 plasmid comprising a codon optimized DNA sequence encoding a HIV pol protein, in
place of the first plasmid comprising a codon optimized DNA sequence encoding both
HIV gagpol proteins.

40 Virus particles produced by the methods described herein, using a codon
optimized HIV packaging construct produced as described herein, were compared by
Western analysis with virus particles produced as described in Naldini *et al., Science,*
25 272:263-267 (1996), using the packaging construct plasmid pCMVΔR8.2. Both the
immunological reactivity and the proteolytic processing were confirmed to be
45 indistinguishable.

50

55

5

-17-

A plasmid comprising a DNA sequence of interest and HIV cis-acting sequences required for packaging, reverse transcription and integration is also referred to herein as a transfer vector. A transfer vector, as used herein, refers to a vehicle which is used to introduce a DNA of interest into a eukaryotic cell, particularly a mammalian cell.

10 5 Figure 3 depicts an example of a transfer vector.

15 DNA sequence of interest, as used herein, include all or a portion of a gene or genes encoding a nucleic acid product whose expression in a cell or a mammal is desired. In a particular embodiment, the nucleic acid product is a heterologous therapeutic protein. Examples of therapeutic proteins include antigens or immunogens, such as a polyclonal vaccine, cytokines, tumor necrosis factor, interferons, interleukins, adenosine deaminase, insulin, T-cell receptors, soluble CD4, growth factors, such as epidermal growth factor, human growth factor, insulin-like growth factors, fibroblast growth factors), blood factors, such as Factor VIII, Factor IX, cytochrome b, glucocerebrosidase, ApoE, ApoC, ApoAI, the LDL receptor, negative selection markers or "suicide proteins", such as thymidine kinase (including the HSV, CMV, VZV TK), anti-angiogenic factors, Fc receptors, plasminogen activators, such as t-PA, u-PA and streptokinase, dopamine, MHC, tumor suppressor genes such as p53 and Rb, monoclonal antibodies or antigen binding fragments thereof, drug resistance genes, ion channels, such as a calcium channel or a potassium channel, adrenergic receptors, hormones (including growth hormones) and anti-cancer agents. In another embodiment, the nucleic acid product is a gene product to be expressed in a cell or a mammal and which product is otherwise defective or absent in the cell or mammal. For example, the nucleic acid product can be a functional gene(s) which is defective or absent in the cell or mammal.

20 25 30 35 40 45 DNA sequence of interest includes DNA sequences (control sequences) which are necessary to drive the expression of the gene or genes. The control sequences are operably linked to the gene. The term "operably linked", as used herein, is defined to mean that the gene is linked to control sequences in a manner which allows expression

50

55

5

-18-

10

of the gene (or the nucleic acid sequence). Generally, operably linked means contiguous.

15

Control sequences include a transcriptional promoter, an optional operator sequence to control transcription, a sequence encoding suitable mRNA ribosomal binding sites and sequences which control termination of transcription and translation. In a particular embodiment, a recombinant gene encoding a desired nucleic acid product can be placed under the regulatory control of a promoter which can be induced or repressed, thereby offering a greater degree of control with respect to the level of the product produced.

20

As used herein, the term "promoter" refers to a sequence of DNA, usually upstream (5') of the coding region of a structural gene, which controls the expression of the coding region by providing recognition and binding sites for RNA polymerase and other factors which may be required for initiation of transcription. Suitable promoters are well known in the art. Exemplary promoters include the SV40, CMV and human elongation factor (EF1) promoters. Other suitable promoters are readily available in the art (see, e.g., Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc., New York (1998); Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 2nd edition, Cold Spring Harbor University Press, New York (1989); and U.S. Patent No. 5,681,735).

25

A DNA sequence of interest can be isolated from nature, modified from native sequences or manufactured *de novo*, as described in, for example, Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons, New York (1998); and Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 2nd edition, Cold Spring Harbor University Press, New York (1989). DNA sequences can be isolated and fused together by methods known in the art, such as exploiting and manufacturing compatible cloning or restriction sites.

30

The packaging cell lines and viral particles of the present invention can be used, *in vitro*, *in vivo* and *ex vivo*, to introduce DNA of interest into a eukaryotic cell (e.g., a

45

50

55

5

-19-

10

mammalian cell) or a mammal (e.g., a human or other mammal or vertebrate). The cells can be obtained commercially or from a depository or obtained directly from a mammal, such as by biopsy. The cells can be obtained from a mammal to whom they will be returned or from another/different mammal of the same or different species. For example, using the packaging cell lines or viral particles of the present invention, DNA of interest can be introduced into nonhuman cells, such as pig cells, which are then introduced into a human. Alternatively, the cell need not be isolated from the mammal where, for example, it is desirable to deliver viral particles of the present invention to the mammal in gene therapy.

15

10 *Ex vivo* therapy has been described, for example, in Kasid *et al.*, *Proc. Natl Acad. Sci. USA*, 87:473 (1990); Rosenberg *et al.*, *N. Engl. J. Med.*, 323:570 (1990); Williams *et al.*, *Nature*, 310:476 (1984); Dick *et al.*, *Cell*, 42:71 (1985); Keller *et al.*, *Nature*, 318:149 (1985); and Anderson *et al.*, United States Patent No. 5,399,346.

20

Methods for administering (introducing) viral particles directly to a mammal are generally known to those practiced in the art. For example, modes of administration include parenteral, injection, mucosal, systemic, implant, intraperitoneal, oral, intradermal, transdermal (e.g., in slow release polymers), intramuscular, intravenous including infusion and/or bolus injection, subcutaneous, topical, epidural, etc. Viral particles of the present invention can, preferably, be administered in a pharmaceutically acceptable carrier, such as saline, sterile water, Ringer's solution, and isotonic sodium chloride solution.

25

40 The dosage of a viral particle of the present invention administered to a mammal, including frequency of administration, will vary depending upon a variety of factors, including mode and route of administration; size, age, sex, health, body weight
25 and diet of the recipient mammal; nature and extent of symptoms of the disease or disorder being treated; kind of concurrent treatment, frequency of treatment, and the effect desired.

30

45

50

55

5

-20-

10

The teachings of all the articles, patents, patent applications and GenBank sequences cited herein are incorporated by reference in their entirety.

15

While this invention has been particularly shown and described with references to preferred embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the spirit and scope of the invention as defined by the appended claims.

20

25

30

35

40

45

50

55

Claims

5

10

15

20

25

30

35

40

45

50

55

5

-21-

CLAIMS

10

What is claimed is:

1. A packaging cell line for producing a viral accessory protein independent HIV-derived retroviral vector particle comprising:
 - 5 a) a mammalian cell;
 - b) a first retroviral nucleotide sequence in the cell which comprises a coding sequence for a HIV *gagpol*, wherein said coding sequence has been mutagenized to improve expression of the HIV gagpol proteins;
 - c) a second retroviral nucleotide sequence in the cell which comprises the coding sequence for a heterologous envelope protein; and
 - 10 d) a third retroviral nucleotide sequence in the cell which comprises a DNA sequence of interest and HIV cis-acting sequences required for packaging, reverse transcription and integration.
2. A packaging cell line of Claim 1 wherein the heterologous envelope protein is the G glycoprotein of vesicular stomatitis virus (VSV G).
3. A packaging cell line of Claim 1 wherein the heterologous envelope protein is the amphotropic envelope of the Moloney leukemia virus.
- 40 4. A packaging cell line of Claim 1 wherein the DNA sequence of interest encodes a heterologous therapeutic protein.
- 45 20 5. A packaging cell line comprising:
 - a) a mammalian cell;

50

55

5

-22-

10

- b) a first retroviral nucleotide sequence in the cell which comprises a coding sequence for a HIV *gagpol*, wherein said coding sequence has been mutagenized to improve expression of the HIV gagpol proteins; and
- c) a second retroviral nucleotide sequence in the cell which comprises a DNA sequence of interest and HIV cis-acting sequences required for packaging, reverse transcription and integration.

15

20

6. A packaging cell line of Claim 5 wherein the DNA sequence of interest encodes a heterologous therapeutic protein.

25

7. A packaging cell line comprising:

30

- a) a mammalian cell;
- b) a first retroviral nucleotide sequence in the cell which comprises a coding sequence for a HIV *gagpol*, wherein said coding sequence has been mutagenized to improve expression of the HIV gagpol proteins; and
- c) a second retroviral nucleotide sequence in the cell which comprises the coding sequence for a heterologous envelope protein.

35

8. A method of producing a packaging cell line for producing a viral accessory protein independent HIV-derived retroviral vector particle, comprising co-transfected mammalian host cells with:

40

- a) a first plasmid comprising a DNA sequence which encodes HIV *gagpol* proteins, wherein said DNA sequence has been mutagenized to improve expression of the HIV gag and pol proteins;
- b) a second plasmid comprising a DNA sequence which encodes a heterologous envelope protein; and

45

50

55

5

-23-

10

- c) a third plasmid comprising a DNA sequence of interest and HIV cis-acting sequences required for packaging, reverse transcription and integration.

15

- 9. A method of Claim 8 wherein the heterologous envelope protein is the G glycoprotein of vesicular stomatitis virus (VSV G).

20

- 10. A method of Claim 8 wherein the heterologous envelope protein is the amphotropic envelope of the Moloney leukemia virus.

25

- 11. A method of Claim 8 wherein the DNA sequence of interest is a heterologous therapeutic protein.

30

- 10 12. A method of producing a viral accessory protein independent HIV-derived retroviral vector particle comprising co-transfected mammalian host cells with:
 - a) a first plasmid comprising a DNA sequence which encodes HIV gagpol proteins, wherein said DNA sequence has been mutagenized to improve expression of the HIV gagpol proteins;

35

- 15 b) a second plasmid comprising a DNA sequence which encodes a heterologous envelope protein; and
 - c) a third plasmid comprising a DNA sequence of interest and HIV cis-acting sequences required for packaging, reverse transcription and integration.

40

- 20 13. A method of Claim 12 wherein the heterologous envelope protein is the G glycoprotein of vesicular stomatitis virus (VSV G).

45

50

55

5

-24-

10 14. A method of Claim 12 wherein the heterologous envelope protein is the
amphotropic envelope of the Moloney leukemia virus.

15 15. A method of Claim 12 wherein the DNA sequence of interest encodes a
heterologous therapeutic protein.

20 5 16. A packaging cell line for producing a viral accessory protein independent
lentivirus-derived retroviral vector particle comprising:
a) a mammalian cell;
b) a first retroviral nucleotide sequence in the cell which comprises a
coding sequence for a lentivirus *gagpol*, wherein said coding sequence
10 has been mutagenized to improve expression of the lentivirus *gagpol*
proteins;
c) a second retroviral nucleotide sequence in the cell which comprises the
coding sequence for a heterologous envelope protein; and
d) a third retroviral nucleotide sequence in the cell which comprises a DNA
sequence of interest and lentivirus *cis*-acting sequences required for
15 packaging, reverse transcription and integration.

35 17. A packaging cell line of Claim 16 wherein the heterologous envelope protein is
the G glycoprotein of vesicular stomatitis virus (VSV G).

40 18. A packaging cell line of Claim 16 wherein the heterologous envelope protein is
20 the amphotropic envelope of the Moloney leukemia virus.

45 19. A packaging cell line of Claim 16 wherein the DNA sequence of interest
encodes a heterologous therapeutic protein.

50

55

5

-25-

10 20. A packaging cell line comprising:

- a) a mammalian cell;
- b) a first retroviral nucleotide sequence in the cell which comprises a coding sequence for lentivirus *gagpol*, wherein said coding sequence has been mutagenized to improve expression of the lentivirus *gagpol* proteins; and
- c) a second retroviral nucleotide sequence in the cell which comprises a DNA sequence of interest and lentivirus *cis*-acting sequences required for packaging, reverse transcription and integration.

15 10 21. A packaging cell line of Claim 20 wherein the DNA sequence of interest encodes a heterologous therapeutic protein.

20 22. A packaging cell line comprising:

- a) a mammalian cell;
- b) a first retroviral nucleotide sequence in the cell which comprises a coding sequence for lentivirus *gagpol*, wherein said coding sequence has been mutagenized to improve expression of the lentivirus *gagpol* proteins; and
- c) a second retroviral nucleotide sequence in the cell which comprises the coding sequence for a heterologous envelope protein.

25 20 23. A method of producing a packaging cell line for producing a viral accessory protein independent lentivirus-derived retroviral vector particle, comprising co-transfected mammalian host cells with:

- a) a first plasmid comprising a DNA sequence which encodes lentivirus *gagpol* proteins, wherein said DNA sequence has been mutagenized to improve expression of the lentivirus *gag* and *pol* proteins;

50

55

5

-26-

10

- b) a second plasmid comprising a DNA sequence which encodes a heterologous envelope protein; and
- c) a third plasmid comprising a DNA sequence of interest and lentivirus cis-acting sequences required for packaging, reverse transcription and integration.

15

- 24. A method of Claim 23 wherein the heterologous envelope protein is the G glycoprotein of vesicular stomatitis virus (VSV G).
- 25. A method of Claim 23 wherein the heterologous envelope protein is the amphotropic envelope of the Moloney leukemia virus.
- 26. A method of Claim 23 wherein the DNA sequence of interest is a heterologous therapeutic protein.
- 27. A method of producing a viral accessory protein independent lentivirus-derived retroviral vector particle comprising co-transfected mammalian host cells with:
 - a first plasmid comprising a DNA sequence which encodes lentivirus *gagpol* proteins, wherein said DNA sequence has been mutagenized to improve expression of the lentivirus *gagpol* proteins;
 - b) a second plasmid comprising a DNA sequence which encodes a heterologous envelope protein; and
 - c) a third plasmid comprising a DNA sequence of interest and lentivirus cis-acting sequences required for packaging, reverse transcription and integration.
- 28. A method of Claim 27 wherein the heterologous envelope protein is the G glycoprotein of vesicular stomatitis virus (VSV G).

45

50

55

5

-27-

10 29. A method of Claim 27 wherein the heterologous envelope protein is the
amphotropic envelope of the Moloney leukemia virus.

15 30. A method of Claim 27 wherein the DNA sequence of interest encodes a
heterologous therapeutic protein.

20 5 31. A viral accessory protein independent HIV-derived retroviral vector particle
produced by the method comprising co-transfected mammalian host cells with:
a) a first plasmid comprising a DNA sequence which encodes HIV *gagpol*
proteins, wherein said DNA sequence has been mutagenized to improve
expression of the HIV *gagpol* proteins;

25 10 b) a second plasmid comprising a DNA sequence which encodes a
heterologous envelope protein; and
c) a third plasmid comprising a DNA sequence of interest and HIV cis-
acting sequences required for packaging, reverse transcription and
integration.

30 15 32. A method of Claim 31 wherein the heterologous envelope protein is the G
glycoprotein of vesicular stomatitis virus (VSV G).

35 33. A method of Claim 31 wherein the heterologous envelope protein is the
amphotropic envelope of the Moloney leukemia virus.

40 40 34. A method of Claim 31 wherein the DNA sequence of interest encodes a
20 heterologous therapeutic protein.

45

50

55

5

-28-

10 35. A viral accessory protein independent lentivirus-derived retroviral vector
particle produced by the method comprising co-transfected mammalian host
cells with:

15 5 a) a first plasmid comprising a DNA sequence which encodes lentivirus—
gagpol proteins, wherein said DNA sequence has been mutagenized to
improve expression of the lentivirus *gagpol* proteins;

20 10 b) a second plasmid comprising a DNA sequence which encodes a
heterologous envelope protein; and

25 15 c) a third plasmid comprising a DNA sequence of interest and lentivirus
cis-acting sequences required for packaging, reverse transcription and
integration.

30 36. A method of Claim 35 wherein the heterologous envelope protein is the G
glycoprotein of vesicular stomatitis virus (VSV G).

35 37. A method of Claim 35 wherein the heterologous envelope protein is the
amphotropic envelope of the Moloney leukemia virus.

40 38. A method of Claim 35 wherein the DNA sequence of interest encodes a
heterologous therapeutic protein.

45 39. Isolated DNA encoding a codon optimized HIV *gagpol*.

40 40. Isolated DNA encoding a codon optimized HIV *gag*.

50 20 41. Isolated DNA of Claim 40 comprising the nucleotide sequence of SEQ ID NO:4.

55 42. Isolated DNA encoding a codon optimized HIV *pol*.

5

-29-

10 43. Isolated DNA of Claim 42 comprising the nucleotide sequence of SEQ ID NO:10.

15 44. A method of introducing a DNA sequence of interest into a mammal comprising introducing into said mammal a viral accessory protein independent HIV-derived retroviral vector particle comprising the DNA sequence of interest.

20 45. The method of Claim 44 wherein the mammal is a human.

25 46. The method of Claim 44 wherein the DNA sequence of interest encodes a heterologous therapeutic protein.

30 47. A method of introducing a DNA sequence of interest into a mammal comprising the steps of:
10 a) introducing into cells a viral accessory protein independent HIV-derived retroviral vector particle comprising the DNA sequence of interest; and
b) returning the cells obtained in step a) to the mammal.

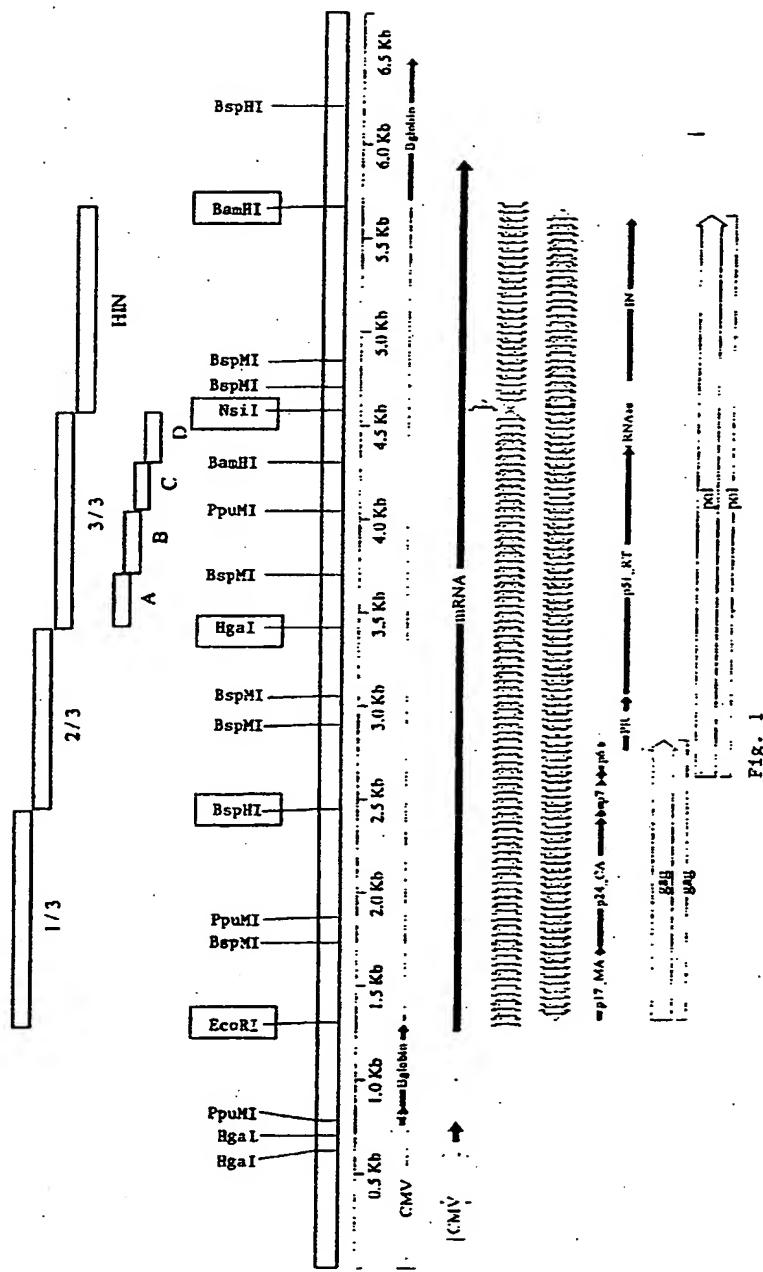
35 48. The method of Claim 47 wherein the mammal is a human.

40 49. The method of Claim 47 wherein the DNA sequence of interest is a heterologous therapeutic protein.

45

50

55



Codon Usage Frequencies

Amino Acid	pNL4-3 gagpol	man gagpol	Amino Acid	pNL4-3 gagpol	man	Amino Acid	pNL4-3 gagpol	man
gca Ala(A)	58	13	ggg Gly(G)	55	14	cca Pro(P)	53	16
gcc Ala(A)	23	53	ggc Gly(G)	12	50	ccc Pro(P)	17	48
gcg Ala(A)	5	17	ggg Gly(G)	27	24	ccg Pro(P)	2	17
gcu Ala(A)	14	17	ggu Gly(G)	6	12	ccu Pro(P)	27	19
aga Arg(R)	63	10	cac His(H)	24	79	agg Ser(S)	29	34
agg Arg(R)	30	18	cau His(H)	76	21	agu Ser(S)	26	10
cga Arg(R)	4	6				uaa Ser(S)	26	5
cgc Arg(R)	0	37				ucc Ser(S)	7	28
cgg Arg(R)	3	21	aua Ile(I)	57	5	ucg Ser(S)	4	9
cgu Arg(R)	0	7	auc Ile(I)	17	77	ucu Ser(S)	6	13
			auu Ile(I)	26	18			
bac Asn(N)	27	78	caa Leu(L)	15	3	aca Thr(T)	52	14
aau Asn(N)	73	22	cuc Leu(L)	10	26	acc Thr(T)	18	57
bac Asp(D)	40	75	cug Leu(L)	11	58	acg Thr(T)	1	15
bau Asp(D)	60	25	ccu Leu(L)	11	5	acu Thr(T)	29	14
ugg Cys(C)	14	68	uuu Leu(L)	40	2	ugg Trp(W)	100	100
ugu Cys(C)	26	32	uug Leu(L)	13	6			
			aaa Lys(K)	69	18	uac Tyr(Y)	26	74
caa Gln(Q)	56	12	aaq Lys(K)	31	82	uaa Tyr(Y)	74	26
cag Gln(Q)	44	88	aug Met(M)	100	100	gua Val(V)	58	5
						guc Val(V)	13	25
gaa Glu(E)	70	25	uuc Phe(F)	40	80	gug Val(V)	16	64
cag Glu(E)	30	75	uuu Phe(F)	60	20	guu Val(V)	14	7

Fig. 2

3/29

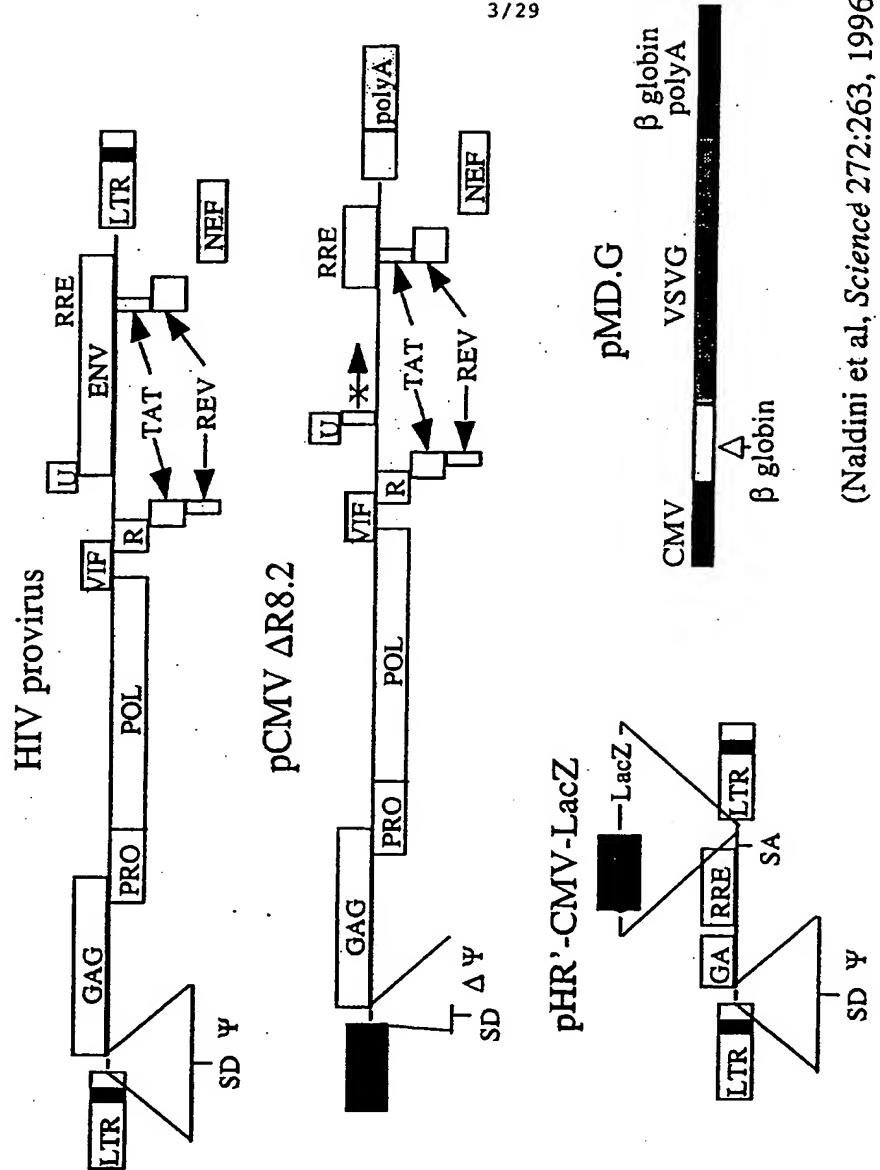


Fig. 3

Rev

- Regulates HIV gene expression by promoting cytoplasmic levels of unspliced and singly spliced mRNAs
- Postulated to affect splicing, stability, transport, and translation

Fig. 4

Codon Optimization of HIV *gagpol*

- Remove A-rich instability elements
- Improve translational efficiency
- Reduce risk of recombination with transfer vector

Fig. 5

6/29

Inactivation of Inhibitory Sequences in *gag*

Schwartz, S., et al.

336

atg ggt gcg aga gcg tca gta tta agc ggg gga gaa tta gat cga tgt gaa aaa att cg
396

M1

tta agg cca ggg gga aag aaa tat aaa tta aaa cat ata gtt gca agg gag
456

G G C GC G C C

cta gaa cga ctt gca gtt atat cct gcc ctg tta gaa aca tca gaa ggc tgt aga caa ata
516

M2

ctg gga cag cta caa cca tcc ctt cag aca gga tca gaa ctt aga tca ttt tat aat
576

M3

M3 G G C C C C C
aca gta gca acc ctc tat tgt gtg cat caa agg ata gag ata aaa gac acc aag gaa gct
636

C GC C C G

M4

tta gac aag ata gag gaa gag caa aac aaa agt aag aaa aaa gca cag caa gca gct
696

GTCC G G C G

gac aca gga cac agc aat cag gtc agc caa aat tac

Fig. 6

7/29

Nucleotide Content of HIV *gagpol*

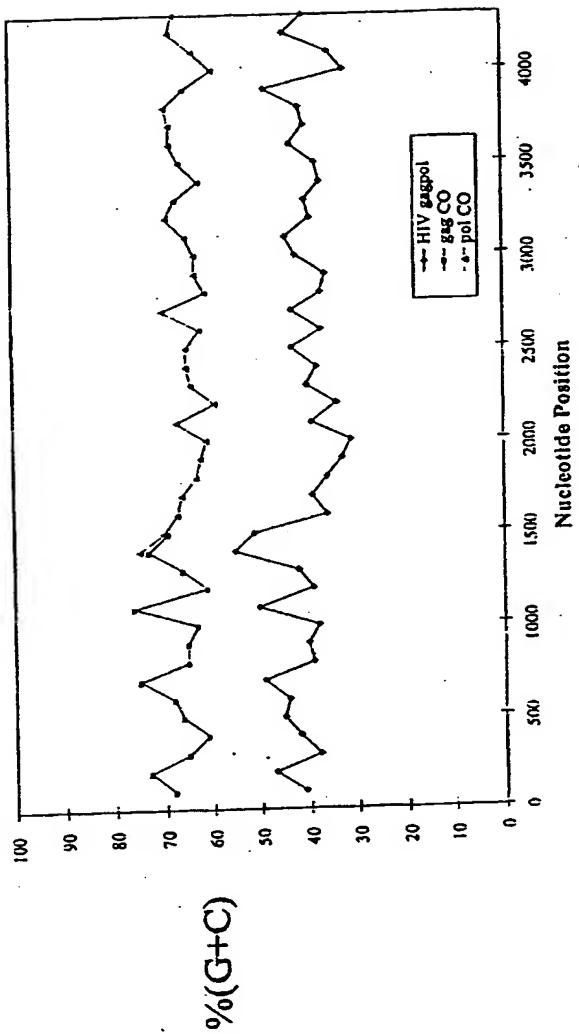


Fig. 7

8/29

Alignment Report of Codon optimization (gag).MEG, using Clustal method with PAM250 residue weight table.

	810	
792	M G A R A S V L S G G E L D K	NL4-3 genbank.seq
792	ATG GGT CGG AGA GCG TCG GTA TTA AGC GGG CGA GAA TTA GAT AAA	
1319	M G A R A S V L S G G E L D K	pHDMHgpm2.seq
1319	ATG GGC GCC CGC GCC TCC GTG CTG TCC GGC GGC GAG CTG GAC AAG	
	840	870
837	W E K I R L R P G G K K Q Y K	NL4-3 genbank.seq
837	TGG GAA AAA ATT CGG TTA AGG CCA CGG GGA AAG AAA CAA TAT AAA	
1364	W E K I R L R P G G K K Q Y K	pHDMHgpm2.seq
1364	TGG GAG AAG ATC CGC CTG CGC CCC GGC AAG AAG CGG TAC AAG	
	900	
882	L K H I V W A S R E L E R F A	NL4-3 genbank.seq
882	CTA AAA CAT ATA GTA TGG GCA AGC AGG GAG CTA GAA CGA TGC TTC GCA	
1409	L K H I V W A S R E L E R F A	pHDMHgpm2.seq
1409	CTG AAG CAC ATC GTG TGG GCC TCC CGC GAG CTG GAG CGC TTC GGC	
	930	960
927	V N P G L L E T S E G C R Q I	NL4-3 genbank.seq
927	GTT AAT CCT CGC CTT TTA GAG ACA TCA GAA GGC TGT AGA CAA ATA	
1454	V N P G L L E T S E G C R Q I	pHDMHgpm2.seq
1454	GTG AAC CCC GGC CTG CTG GAG ACC TCC GAG GGC TGC CGC CAG ATC	
	990	
972	L G Q L Q P S L Q T G S E E L	NL4-3 genbank.seq
972	CTG GGA CAG CTA CAA CCA TCC CTT CAG ACA GGA TCA GAA GAA CTT	
1499	L G Q L Q P S L Q T G S E E L	pHDMHgpm2.seq
1499	CTG GGC CAG CTG CAG CCC TCC CTG CAA ACC GGC TCC GAG GAG CTG	
	1020	1053
1017	R S L Y N T I A V L Y C V H Q	NL4-3 genbank.seq
1017	AGA TCA TTA TAT AAT ACA ATA GCA GTC CTC TAT TGT GTG CAT CAA	
1544	R S L Y N T I A V L Y C V H Q	pHDMHgpm2.seq
1544	CGC TCC CTG TAC AAC ACC ATC GGC GTG CTG TAC TGC GTG CAC CAG	
	1080	
1062	R I D V K D T K E A L D K I E	NL4-3 genbank.seq
1062	AGG ATA GAT GTA AAA GAC ACC AAG GAA GCC TTA GAT AAG ATA GAG	
1589	R I D V K D T K E A L D K I E	pHDMHgpm2.seq
1589	CGC ATC GAC GTG AAG GAC ACC AAG GAG GCC CTG GAC AAG ATC GAG	
	1110	1140
1107	E E Q N K S K K K A Q Q A A A	NL4-3 genbank.seq
1107	GAA GAG CAA AAC AAA AGT AAG AAA AAG GCA CAG CAA GCA GCA GCT	
1634	E E Q N K S K K K A Q Q A A A	pHDMHgpm2.seq
1634	GAG GAG CAG AAC AAG TCC AAG AAG AAG GGC CAG CAG GCC GCC GGC	

Fig. 8A

9/29

Alignment Report of Codon optimization (gag).MEG, using Clustal method with PAM250 residue weight table.

1170																					
1152	D	T	G	N	N	S	'Q	V	S	Q	N	Y	P	I	V	NL4-3 genbank.SEQ					
1152	GAC	ACA	GGG	AAC	AAC	AGC	AGC	CAG	GTC	AGC	CAA	AAT	TAC	CCT	ATA	GTG					
1679	D	T	G	N	N	S	'Q	V	S	Q	N	Y	P	I	V	pHDMHgpm2.seq					
1679	GAC	ACC	GGC	AAC	AAC	TCC	CAG	GTG	TCC	CAG	AAC	TAC	CCC	ATC	GTG						
1200																					
1197	Q	N	L	Q	G	Q	M	V	H	Q	A	I	S	P	R	NL4-3 genbank.SEQ					
1197	CAG	AAC	CTC	CAG	GGG	CAA	ATG	GTA	CAT	CAG	GCC	ATA	TCA	CCT	AGA						
1724	Q	N	L	Q	G	Q	M	V	H	Q	A	I	S	P	R	pHDMHgpm2.seq					
1724	CAG	AAC	CTG	CAG	GGC	CAG	ATG	GTG	CAC	CAG	GCC	ATC	TCC	CCC	CGC						
1260																					
1242	T	L	N	A	W	V	K	V	V	E	E	K	A	F	S	NL4-3 genbank.SEQ					
1242	ACT	TTA	ATT	GCA	TGG	GTA	AAA	GTA	GAA	GAG	AAG	GCT	TTC	AGC							
1769	T	L	N	A	W	V	K	V	V	E	E	K	A	F	S	pHDMHgpm2.seq					
1769	ACC	CTG	AAC	GCC	TGG	GTG	AAG	GTG	GTG	GAG	GAG	AAG	GCC	TTC	TCC						
1290																					
1287	P	E	V	I	P	M	F	S	A	L	S	E	G	A	T	NL4-3 genbank.SEQ					
1287	CCA	GAA	GTA	ATA	CCC	ATG	TTT	TCA	GCA	TAA	TCA	GAA	GGA	GCC	ACC						
1814	P	E	V	I	P	M	F	S	A	L	S	Z	G	A	T	pHDMHgpm2.seq					
1814	CCC	GAA	GTC	ATC	CCC	ATG	TTC	TCC	GCC	CTG	TCC	GAG	GGC	GCC	ACC						
1320																					
1327	P	Q	D	L	N	T	M	L	N	T	V	G	G	H	Q	NL4-3 genbank.SEQ					
1327	CCA	CAA	GAT	TTA	ATT	ACC	ATG	CTA	AAC	ACA	GTG	GGG	GGG	CAT	CAA						
1859	P	Q	D	L	N	T	M	L	N	T	V	G	G	H	Q	pHDMHgpm2.seq					
1859	CCC	CAG	GAC	CTG	AAC	ACC	ATG	CTG	AAC	ACC	GTG	GGG	GGC	GCC	CAC						
1350																					
1377	A	A	M	Q	M	L	K	E	T	I	N	E	S	A	A	NL4-3 genbank.SEQ					
1377	GCA	GCC	ATG	CAA	ATG	TTA	AAA	GAG	ACC	ATC	ATT	GAG	GAA	GCT	GCA						
1904	A	A	M	Q	M	L	K	E	T	I	N	E	E	A	A	pHDMHgpm2.seq					
1904	GCC	GCC	ATG	CAG	ATG	CTG	AAG	GAG	ACC	ATC	AAC	GAG	GAG	GCC	GCC						
1380																					
1377	A	A	M	Q	M	L	K	E	T	I	N	E	S	A	A	NL4-3 genbank.SEQ					
1377	GCA	GCC	ATG	CAA	ATG	TTA	AAA	GAG	ACC	ATC	ATT	GAG	GAA	GCT	GCA						
1904	A	A	M	Q	M	L	K	E	T	I	N	E	E	A	A	pHDMHgpm2.seq					
1904	GCC	GCC	ATG	CAG	ATG	CTG	AAG	GAG	ACC	ATC	AAC	GAG	GAG	GCC	GCC						
1410																					
1422	E	W	D	R	L	H	P	V	B	A	G	P	I	A	P	NL4-3 genbank.SEQ					
1422	GAA	TGG	GAT	AGA	TTG	CAT	CCA	GTG	CAT	GCA	GGG	CCT	ATT	GCA	CCA						
1949	E	W	D	R	L	H	P	V	B	A	G	P	I	A	P	pHDMHgpm2.seq					
1949	GAG	TGG	GAC	CGC	CTG	CAC	CCC	GTG	CAC	GCC	GGC	CCC	ATC	GCC	CCC						
1440																					
1467	G	Q	M	R	E	P	R	G	S	D	I	A	G	T	T	NL4-3 genbank.SEQ					
1467	GGC	CAG	ATG	AGA	GAA	CCA	AGG	GGG	AGT	GAC	ATA	GCA	GGG	ACT	ACT						
1994	G	Q	M	R	E	P	R	G	S	D	I	A	G	T	T	pHDMEgpm2.seq					
1994	GGC	CAG	ATG	CGC	GAG	CCC	CGC	GGC	TCC	GAC	ATC	GCC	GGC	ACC	ACC						
1470																					
1470																					

Fig. 8B

10/29

Alignment Report of Codon optimization (gag).MEG, using Clustal method with PAM250 residue weight table.

1530															
1512	S	T	L	Q	E	Q	I	G	W	M	T	H	N	P	P
1512	AGT	ACC	CTT	CAG	GAA	CAA	ATA	GGA	TGG	ATG	ACA	CAT	AAT	CCA	CCT
2039	S	T	L	Q	E	Q	I	G	W	M	T	H	N	P	P
2039	TCC	ACC	CTG	CAA	GAG	CAG	ATC	GGC	TGG	ATG	ACC	CAC	AAC	CCC	CCC
1560															
1557	I	P	V	G	E	I	Y	K	R	W	I	I	L	G	L
1557	ATC	CCA	GTA	GGG	GAA	ATC	TAT	AAA	AGA	TGG	ATA	ATC	CTG	GGA	TTA
2084	I	P	V	G	E	I	Y	K	R	W	I	I	L	G	L
2084	ATC	CCC	GTG	GGC	GAG	ATC	TAC	AAG	GGC	TGG	ATC	ATC	CTG	GGC	CTG
1620															
1602	N	K	I	V	R	M	Y	S	P	T	S	I	L	D	I
1602	ATC	AAA	ATA	GTA	AGA	ATG	TAT	AGC	CCT	ACC	AGC	ATT	CTG	GAC	ATA
2129	N	K	I	V	R	M	Y	S	P	T	S	I	L	D	I
2129	AAC	AAG	ATC	GTG	GGC	ATG	TAC	TCC	CCC	ACC	TCC	ATC	CTG	GAC	ATC
1650															
1647	R	Q	G	P	K	E	P	F	R	D	Y	V	D	R	F
1647	AGA	CAA	GGG	CCA	AAG	GAA	CCC	TTT	AGA	GAC	TAT	GTA	GAC	CGA	TTC
2174	R	Q	G	P	K	E	P	F	R	D	Y	V	D	R	F
2174	CGC	CAG	GGC	CCC	AAG	GAG	CCC	TTC	CGC	GAC	TAC	GTG	GAC	CGC	TTC
1710															
1692	Y	K	T	L	R	A	E	Q	A	S	Q	E	V	K	N
1692	TAT	AAA	ACT	CTA	AGA	GCC	GAG	CAA	GCT	TCA	CAA	GAG	GTA	AAA	AAT
2219	Y	K	T	L	R	A	E	Q	A	S	Q	E	V	K	N
2219	TAC	AAG	ACC	CTG	CGC	GCC	GAG	CAG	GCC	TCC	CAG	GAG	GTA	AAG	AAC
1740															
1737	W	M	T	E	T	L	L	V	Q	N	A	N	P	D	C
1737	TGG	ATG	ACA	GAA	ACC	TTG	TTG	GTC	CAA	ATT	CGG	AAC	CCA	GAT	TGT
2264	W	M	T	E	T	L	L	V	Q	N	A	N	P	D	C
2264	TGG	ATG	ACC	GAG	ACC	CTG	CTG	GTG	CAG	AAC	GCC	AAC	CCC	GAC	TGC
1800															
1782	K	T	I	L	K	A	L	G	P	G	A	T	L	E	E
1782	AAG	ACT	ATT	TTA	AAA	GCA	TTG	GGG	CCA	GGG	ACA	ACA	CTA	GAA	GAA
2309	K	T	I	L	K	A	L	G	P	G	A	T	L	E	E
2309	AAG	ACC	ATC	CTG	AAG	GCC	CTG	GGC	CCC	GGC	GCC	ACC	CTG	GAG	GAG
1830															
1827	M	M	T	A	C	Q	G	V	G	G	P	G	H	K	A
1827	ATG	ATG	ACA	GCA	TGT	CAG	GGG	GGG	GGG	CCC	GGC	CAT	AAA	GCA	
2354	M	M	T	A	C	Q	G	V	G	G	P	G	H	K	A
2354	ATG	ATG	ACC	GCC	TGC	CAG	GGC	GTG	GGC	GGC	CCC	GGC	CAC	AAG	GCC
1860															
1827	M	M	T	A	C	Q	G	V	G	G	P	G	H	K	A
1827	ATG	ATG	ACA	GCA	TGT	CAG	GGG	GGG	GGG	CCC	GGC	CAT	AAA	GCA	
2354	M	M	T	A	C	Q	G	V	G	G	P	G	H	K	A
2354	ATG	ATG	ACC	GCC	TGC	CAG	GGC	GTG	GGC	GGC	CCC	GGC	CAC	AAG	GCC

Fig. 8C

11/29

Alignment Report of Codon optimization (gag).MEG, using Clustal method with PAM250 residue weight table.

1890																	
1872	R	V	L	A	E	A	M	S	Q	V	T	N	P	A	T		
1872	AGA	GTT	TTG	GCT	GAA	GCA	ATG	AGC	CAA	GTA	ACA	AAT	CCA	GCT	ACC	NL4-3	genbank.SEQ
2399	R	V	L	A	E	A	M	S	Q	V	T	N	P	A	T		
2399	CGC	GTG	CTG	GCC	GAG	GCC	ATG	TCC	CAA	GTC	ACC	AAC	CCC	GCC	ACC	pHDMEgpm2.seq	
1920																	
1917	I	M	I	Q	K	G	N	F	R	N	Q	R	R	K	T	V	
1917	ATA	ATG	ATA	CAG	AAA	GCC	AAT	TTT	AGG	ACG	CAA	AGA	AAG	ACT	GTT	NL4-3	genbank.SEQ
2444	I	M	I	Q	K	G	N	F	R	N	Q	R	K	T	V		
2444	ATC	ATG	ATC	CAG	AAG	GCC	AAC	TTC	CGC	AAC	CGG	CGC	AAG	ACC	GTG	pHDMEgpm2.seq	
1980																	
1962	K	C	F	N	C	G	K	E	G	H	I	A	K	N	C		
1962	AAG	TGT	TTC	AAT	TGT	GCC	AAA	GAA	GGG	CAC	ATA	GCC	AAA	ATT	TGC	NL4-3	genbank.SEQ
2489	K	C	F	N	C	G	K	E	G	H	I	A	K	N	C		
2489	AAG	TGC	TTC	AAC	TGC	GCC	AAG	GAG	GGG	CAC	ATC	GCC	AAG	AAC	TGC	pHDMEgpm2.seq	
2010																	
2007	R	A	P	R	K	K	G	C	W	K	C	G	K	E	G		
2007	AGG	GCC	CCT	AGG	AAA	AAG	GGC	TGT	TGG	AAA	TGT	GGA	AAG	GAA	GGG	NL4-3	genbank.SEQ
2534	R	A	P	R	K	K	G	C	W	K	C	G	K	Z	G		
2534	CGC	GCC	CCC	CGC	AAG	AAG	GGC	TGC	TGG	AAG	TGC	GCC	AAG	GAG	GGC	pHDMEgpm2.seq	
2070																	
2052	H	Q	M	K	D	C	T	Z	R	Q	A	N	F	L	G		
2052	CAC	CAA	ATG	AAA	GAT	TGT	ACT	GAG	AGA	CAG	GCT	AAT	TTT	TAA	GGG	NL4-3	genbank.SEQ
2579	H	Q	M	K	D	C	T	E	R	Q	A	N	F	L	G		
2579	CAC	CAG	ATG	AAA	GAT	TGT	ACT	GAG	AGA	CAG	GCT	AAT	TTT	TAA	GGG	pHDMEgpm2.seq	
2100																	
2097	K	I	N	P	S	H	K	G	R	P	G	N	F	Z	Q		
2097	AAG	ATC	TGG	CCT	TCC	CAC	AAG	GGA	AGG	CCA	GGG	AAT	TTT	CTT	CAG	NL4-3	genbank.SEQ
2624	K	I	W	P	S	H	K	G	R	P	G	N	F	L	Q		
2624	AAG	ATC	TGG	CCT	TCC	CAC	AAG	GGA	AGG	CCA	GGG	AAT	TTT	CTT	CAG	pHDMEgpm2.seq	
2160																	
2142	S	R	P	E	P	T	A	P	P	E	E	S	F	R	E		
2142	AGC	AGA	CCA	GAG	CCA	ACA	GCA	GCC	CCA	GAA	GAG	AGC	TTC	AGG	TTT	NL4-3	genbank.SEQ
2669	S	R	P	E	P	T	A	P	P	E	E	S	F	R	E		
2669	AGC	AGA	CCA	GAG	CCA	ACA	GCA	GCC	CCA	GAA	GAG	AGC	TTC	AGG	TTT	pHDMEgpm2.seq	
2190																	
2187	G	E	E	T	T	T	P	S	Q	K	Q	Z	P	I	D		
2187	GGG	GAA	GAG	ACA	ACA	ACT	CCC	TCT	CAG	AAG	CAG	GAG	CCG	ATA	GAC	NL4-3	genbank.SEQ
2714	G	E	E	T	T	T	P	S	Q	K	Q	Z	P	I	D		
2714	GGG	GAA	GAG	ACA	ACA	ACT	CCC	TCT	CAG	AAG	CAG	GAG	CCG	ATA	GAC	pHDMEgpm2.seq	
2220																	

Fig. 8D

12/29

Alignment Report of Codon optimization (gag) MEG, using Clustal method with PAM250 residue weight table.

															NL4-3 genbank. SEQ	
2232	R	E	L	Y	P	L	A	S	L	R	S	L	F	G	S	NL4-3 genbank. SEQ
2232	AAG	GAA	CTG	TAT	CCT	TTA	GCT	TCC	CTC	AGA	TCA	CTC	TTT	GGC	AGC	
2759	R	E	L	Y	P	L	A	S	L	R	S	L	F	G	S	pHDMHgpm2.seq
2759	AAG	GAA	CTG	TAT	CCT	TTA	GCT	TCC	CTC	AGA	TCA	CTC	TTT	GGC	AGC	
2280																
2277	D	P	S	S	S	S	Q									NL4-3 genbank. SEQ
2277	GAC	CCC	TCG	TCA	CRA	TTA										
2804	D	P	S	S	S	S	Q									pHDMHgpm2.seq
2804	GAC	CCC	TCG	TCA	CRA	TTA										

Fig. 8E

Alignment Report of Codon Optimization (pol) MEG, using Clustal method with PAM250 residue weight table.

	2090	2129	
2087	F F R E D L A F P Q G X A R E	NL4-3 genbank.seq	
2087	TTT TTT AGG GAA GAT CTG GCC TTC CCA CAA GGG AAG GCC AGG GAA		
2085	F F R E D L A F P Q G K A R E	pNL4-3.seq	
2085	TTT TTT AGG GAA GAT CTG GCC TTC CCA CAA GGG AAG GCC AGG GAA		
2612	F F R E D L A F P Q G K A R E	pHDMHgpm2.seq	
2612	TTT TTT AGG GAA GAT CTG GCC TTC CCA CAA GGG AAG GCC AGG GAA		
	2150		
2132	F S S E Q T R A N S P T R R E	NL4-3 genbank.seq	
2132	TTC TCT TCA GAG CAG ACC AGA GCC AAC AGC CCC ACC AGA AGA GAG		
2130	F S S E Q T R A N S P T R R E	pNL4-3.seq	
2130	TTC TCT TCA GAG CAG ACC AGA GCC AAC AGC CCC ACC AGA AGA GAG		
2657	F S S E Q T R A N S P T R R E	pHDMHgpm2.seq	
2657	TTC TCT TCA GAG CAG ACC AGA GCC AAC AGC CCC ACC AGA AGA GAG		
	2180	2210	
2177	L Q V W G R D N N S L S E A A G	NL4-3 genbank.seq	
2177	CTT CAG GTT TGG GGA AGA GAC AAC TCC CTC TCA GAA GCA GGA		
2175	L Q V W G R D N N S L S E A A G	pNL4-3.seq	
2175	CTT CAG GTT TGG GGA AGA GAC AAC TAC TCC CTC TCA GAA GCA GGA		
2702	L Q V W G R D N N S L S E A A G	pHDMHgpm2.seq	
2702	CTT CAG GTT TGG GGA AGA GAC AAC TCC CTC TCA GAA GCA GGA		
	2240		
2222	A D R Q G T V S F S E P Q I T	NL4-3 genbank.seq	
2222	GCC GAT AGA CAA GGA ACT GTA TCC TTT AGC TTC CCT CAG ATC ACT		
2220	A D R Q G T V S F S F P Q I T	pNL4-3.seq	
2220	GCC GAT AGA CAA GGA ACT GTA TCC TTT AGC TTC CCT CAG ATC ACT		
2747	A D R Q G T V S F E S F P Q I T	pHDMHgpm2.seq	
2747	GCC GAT AGA CAA GGA ACT GTA TCC TTT AGC TTC CCT CAG ATC ACT		
	2270	2300	
2267	L W Q R P L V T I K I G G Q L	NL4-3 genbank.seq	
2267	CTT TGG CAG CGA CCC CTC GTC ACA ATA AAG ATA GGG GGG CAA TTA		
2265	L W Q R P L V T I K I G G Q L	pNL4-3.seq	
2265	CTT TGG CAG CGA CCC CTC GTC ACA ATA AAG ATA GGG GGG CAA TTA		
2792	L W Q R P L V T I K I G G Q L	pHDMHgpm2.seq	
2792	CTT TGG CAG CGA CCC CTC GTC ACA ATA AAG ATC GGT GGC CAG CTG		
	2330		
2312	K E A L L D T G A D D T V L E	NL4-3 genbank.seq	
2312	AAG GAA GCT CTA TTA GAT ACA GGA GCA GAT GAT ACA GTC TTA GAA		
2310	K E A L L D T G A D D T V L E	pNL4-3.seq	
2310	AAG GAA GCT CTA TTA GAT ACA GGA GCA GAT GAT ACA GTC TTA GAA		
2837	K E A L L D T G A D D T V L E	pHDMHgpm2.seq	
2837	AAG GAG GCC CTG GAC ACC GGC GCC GAC ACC GTG CTG GAG		

Fig. 9A

14/29

Alignment Report of Codon Optimization (pol).MEG, using Clustal method with PAM250 residue weight table.

	2360	2390	
2357	E M N L P G R W K P K M I G G		NL4-3 genbank.seq
2357	GAA ATG AAT TTG CCA GGA AGA TGG AAA CCA AAA ATG ATA GGG GGA		
2355	E M N L P G R W K P K M I G G		pNL4-3.seq
2355	GAA ATG AAT TTG CCA GGA AGA TGG AAA CCA AAA ATG ATA GGG GGA		
2882	E M N L P G R W K P K M I G G		pHDMHgpm2.seq
2882	GAG ATG AAC CTG CCC GGC CGC TGG AAG CCC AAG ATG ATC GCC GGC		
	2420		
2402	I G G F I R V G Q Y D Q I L I		NL4-3 genbank.seq
2402	ATT GGA GGT TTT ATC AAA GTA GGA CAG TAT GAT CAG ATA CTC ATA		
2400	I G G F I K V R Q Y D Q I L I		pNL4-3.seq
2400	ATT GGA GGT TTT ATC AAA GTA AGA CAG TAT GAT CAG ATA CTC ATA		
2927	I G G F I K V R Q Y D Q I L I		pHDMHgpm2.seq
2927	ATC GGC GGC TTC ATC AAA GTC CGC CAG TAC GAC CAG ATC CTG ATC		
	2450	2480	
2447	E I C G H K A I G T V L V G P		NL4-3 genbank.seq
2447	GAA ATC TGC GGA CAT AAA GCT ATA GGT ACA GCA TTA GTA GGA CCT		
2445	E I C G H K A I G T V L V G P		pNL4-3.seq
2445	GAA ATC TGC GGA CAT AAA GCT ATA GGT ACA GCA TTA GTA GGA CCT		
2972	E I C G H K A I G T V L V G P		pHDMHgpm2.seq
2972	GAG ATC TGC GGC CAC AAG GCC ATC GGC ACC GTG CTG GTG GGC CCC		
	2510		
2492	T P V N I I G R N L L T Q I G		NL4-3 genbank.seq
2492	ACA CCT GTC AAC ATA ATT GGA AGA AAT CTG TTG ACT CAG ATT GGC		
2490	T P V N I I G R N L L T Q I G		pNL4-3.seq
2490	ACA CCT GTC AAC ATA ATT GGA AGA AAT CTG TTG ACT CAG ATT GGC		
3017	T P V N I I G R N L L T Q I G		pHDMHgpm2.seq
3017	ACC CCC GTG AAC ATC ATC GGC CGC AAC CTG ACC CAG ATC GGC		
	2540	2570	
2537	C T L N F P I S P I E T V P V		NL4-3 genbank.seq
2537	TGC ACT TTA ATT TTT CCC ATT AGT CCT ATT GAG ACT GCA CCA GTC		
2535	C T L N F P I S P I E T V P V		pNL4-3.seq
2535	TGC ACT TTA ATT TTT CCC ATT AGT CCT ATT GAG ACT GCA CCA GTC		
3062	C T L N F P I S P I E T V P V		pHDMHgpm2.seq
3062	TGC ACC CTG AAC TTC CCC ATC TCC CCC ATC GAG ACC GTG CCC GTG		
	2600		
2582	K L K P G M D G P K V K Q W P		NL4-3 genbank.seq
2582	AAA TTA AAG CCA GGA ATG GAT GGC CCA AAA GTT AAA CAA TGG CCA		
2580	K L K P G M D G P K V K Q W P		pNL4-3.seq
2580	AAA TTA AAG CCA GGA ATG GAT GGC CCA AAA GTT AAA CAA TGG CCA		
3107	K L K P G M D G P K V K Q W P		pHDMHgpm2.seq
3107	AAG CTG AAG CCC GGC ATG GAC GGC CCC AAA GTC AAG CAG TGG CCC		

Fig. 9B

15/29

Alignment Report of Codon Optimization (pol)MEG, using Clustal method with PAM250 residue weight table.

	2630	2660	
2627	L T E S K I K A L V E I C T E		NL4-3 genbank.SEQ
2627	TTC ACA GAA GAA AAA ATA AAA GCA TTA GTA GAA ATT TGT ACA GAA		
2625	L T E E K I K A L V E I C T E		pNL4-3.seq
2625	TTC ACA GAA GAA AAA ATA AAA GCA TTA GTA GAA ATT TGT ACA GAA		
3152	L T E E K I K A L V E I C T E		pHDMHgpm2.seq
3152	CTG ACC GAG GAG ATC AAG GCC CTG GTG GAG ATC TGC ACC GAG		
	2690		
2672	M E K E G K I S . K I G P E N P		NL4-3 genbank.SEQ
2672	ATG GAA AAG GAA GGA AAA ATT TCA AAA ATT GGG CCT GAA AAT CCA		
2670	M E K E G K I S . K I G P E N P		pNL4-3.seq
2670	ATG GAA AAG GAA GGA AAA ATT TCA AAA ATT GGG CCT GAA AAT CCA		
3197	M E K E G K I S . K I G P E N P		pHDMHgpm2.seq
3197	ATG GAG AAG GAG GCC AAG ATC TCC AAG ATC TGC CCC GAG AAC CCC		
	2720	2750	
2717	Y N T P V F A I K K K D S T K		NL4-3 genbank.SEQ
2717	TAC AAT ACT CCA GTA TTT GCC ATA AAG AAA GAC AGT ACT AAA		
2715	Y N T P V F A I K K K D S T K		pNL4-3.seq
2715	TAC AAT ACT CCA GTA TTT GCC ATA AAG AAA GAC AGT ACT AAA		
3242	Y N T P V F A I K K K D S T K		pHDMHgpm2.seq
3242	TAC AAC ACC CCC GTG TTC GCC ATC AAG AAG AAG GAC TCC ACC AAG		
	2780		
2762	W R K L V D F R E L S K R T Q		NL4-3 genbank.SEQ
2762	TGG AGA AAA TTA GTA GAT TTC AGA GAA CTT ATT AAG AGA ACT CAA		
2760	W R K L V D F R E L N K R T Q		pNL4-3.seq
2760	TGG AGA AAA TTA GTA GAT TTC AGA GAA CTT ATT AAG AGA ACT CAA		
3287	W R K L V D F R E L N K R T Q		pHDMHgpm2.seq
3297	TGG CGC AAG CTG GTG GAC TTC CSC GAG CTG AAC AAG CGC ACC CAG		
	2810	2940	
2907	D F W E V Q L G I P H P A G L		NL4-3 genbank.SEQ
2807	GAT TTC TGG GAA GTT CAA TTA GGA ATA CCA CAT CCT GCA GGG TTA		
2805	D F W E V Q L G I P H P A G L		pNL4-3.seq
2805	GAT TTC TGG GAA GTT CAA TTA GGA ATA CCA CAT CCT GCA GGG TTA		
3332	D F W E V Q L G I P H P A G L		pHDMHgpm2.seq
3332	GAC TTC TGG GAG GTG CAG CTG GCC ATC CCC CAC CCC GCC GGC CTG		
	2870		
2852	K Q K K S V T V L D V G D A Y		NL4-3 genbank.SEQ
2852	AAA CAG AAA AAA TCA GTA ACA GCA CTG GAT GTG GGC GAT GCA TAT		
2850	K Q K K S V T V L D V G D A Y		pNL4-3.seq
2850	AAA CAG AAA AAA TCA GTA ACA GCA CTG GAT GTG GGC GAT GCA TAT		
3377	K Q K K S V T V L D V G D A Y		pHDMHgpm2.seq
3377	AAG CAG AAG AAG TCC GTG ACC GTC CTG GAC GTG GGC GAC GCG TAC		

Fig. 9C

16/29

Alignment Report of Codon Optimization (pol)MEG, using Clustal method with PAM250 residue weight table.

	2900														NL4-3 genbank.SEQ	
	F	S	V	P	L	D	K	D	F	R	K	Y	T	A	F	NL4-3 genbank.SEQ
2897	TTT	TCA	GTT	CCC	TTA	GAT	AAA	GAC	TTC	AGG	AGG	TAT	ACT	GCA	TTT	pNL4-3.seq
2895	F	S	V	P	L	D	K	D	F	R	K	Y	T	A	F	pHDMHgpm2.seq
3422	TTC	TCC	GTG	CCC	CTG	GAC	AAG	GAC	TTC	AGG	AGG	TAT	ACT	GCA	TTT	pHDMHgpm2.seq
3422	TTC	TCC	GTG	CCC	CTG	GAC	AAG	GAC	TTC	CGC	AGG	TAC	ACC	GCC	TTC	
	2960														NL4-3 genbank.SEQ	
	T	I	P	S	I	N	N	E	T	P	G	I	R	Y	Q	NL4-3 genbank.SEQ
2942	ACC	ATA	CCT	AGT	ATA	AAC	AAT	GAG	ACA	CCA	GGG	ATT	AGA	TAT	CAG	pNL4-3.seq
2940	T	I	P	S	I	N	N	E	T	P	G	I	R	Y	Q	pHDMHgpm2.seq
3467	ACC	ATA	CCT	AGT	ATA	AAC	AAT	GAG	ACA	CCA	GGG	ATT	AGA	TAT	CAG	
3467	ACC	ATC	CCC	TCC	ATC	AAC	AAC	GAG	ACC	CCC	GGC	ATC	CGC	TAC	CAG	
	2990														3020	NL4-3 genbank.SEQ
	Y	N	V	L	P	Q	G	W	K	G	S	P	A	I	F	NL4-3 genbank.SEQ
2987	TAC	AAT	GTG	CTT	CCA	CGG	GGA	TGG	AAA	GGA	TCA	CCA	GCA	ATA	TTC	pNL4-3.seq
2985	Y	N	V	L	P	Q	G	W	K	G	S	P	A	I	F	pHDMHgpm2.seq
3512	TAC	AAT	GTG	CTT	CCA	CGG	GGA	TGG	AAA	GGA	TCA	CCA	GCA	ATA	TTC	
3512	Y	N	V	L	P	Q	G	W	K	G	S	P	A	I	F	
	3050														NL4-3 genbank.SEQ	
	Q	C	S	M	T	K	I	L	E	P	F	R	K	Q	N	NL4-3 genbank.SEQ
3032	CAG	TGT	AGC	ATG	ACA	AAA	ATC	TTA	GAG	CCT	TTT	AGA	AAA	CAA	AAT	pNL4-3.seq
3030	Q	C	S	M	T	K	I	L	E	P	F	R	K	Q	N	pHDMHgpm2.seq
3557	CAG	TGT	AGC	ATG	ACA	AAA	ATC	TTA	GAG	CCT	TTT	AGA	AAA	CAA	AAT	
3557	Q	C	S	M	T	K	I	L	E	P	F	R	K	Q	N	
	3080														3110	NL4-3 genbank.SEQ
	P	D	I	V	I	Y	Q	Y	M	D	D	L	Y	V	G	NL4-3 genbank.SEQ
3077	CCA	GAC	ATA	GTC	ATC	TAT	CAA	TAC	ATG	GAT	GAT	TTG	TAT	GTA	GGA	pNL4-3.seq
3075	P	D	I	V	I	Y	Q	Y	M	D	D	L	Y	V	G	pHDMHgpm2.seq
3602	CCA	GAC	ATA	GTC	ATC	TAT	CAA	TAC	ATG	GAT	GAT	TTG	TAT	GTA	GGA	
3602	P	D	I	V	I	Y	Q	Y	M	D	D	L	Y	V	G	
	3140														NL4-3 genbank.SEQ	
	S	D	L	E	I	G	Q	H	R	T	K	I	E	E	L	NL4-3 genbank.SEQ
3122	TCT	GAC	TTA	GAA	ATA	GGG	CAG	CAT	AGA	ACA	AAA	ATA	GAG	GAA	CTG	pNL4-3.seq
3120	S	D	L	E	I	G	Q	H	R	T	K	I	E	E	L	pHDMHgpm2.seq
3647	TCT	GAC	TTA	GAA	ATA	GGG	CAG	CAT	AGA	ACA	AAA	ATA	GAG	GAA	CTG	
3647	S	D	L	E	I	G	Q	H	R	T	K	I	E	E	L	

Fig. 9D

17/29

Alignment Report of Codon Optimization (pol)MEG, using Clustal method with PAM250 residue weight table.

	3170	3200	
3167	R Q H L L R W G F T T P D K K	NL4-3 genbank.seq	
3167	AGA CAA CAT CTG TTG AGG TGG GGA TTT ACC ACA CCA GAC AAA AAA		
3165	R Q H L L R W G F T T P D K K	pNL4-3.seq	
3165	AGA CAA CAT CTG TTG AGG TGG GGA TTT ACC ACA CCA GAC AAA AAA		
3692	R Q H L L R W G F T T P D K K	PHDMHgpm2.seq	
3692	CCC CAG CAC CTG CTG CGC TGG GGC TTC ACC ACC CCC GAC AAG AAG		
<hr/>			
	3230		
3212	H Q K E P P F L W M G Y E L H	NL4-3 genbank.seq	
3212	CAT CAG AAA GAA CCT CCA TTC CTT TGG ATG GGT TAT GAA CTC CAT		
3210	H Q K E P P F L W M G Y E L H	pNL4-3.seq	
3210	CAT CAG AAA GAA CCT CCA TTC CTT TGG ATG GGT TAT GAA CTC CAT		
3737	H Q K E P P F L W M G Y E L H	PHDMHgpm2.seq	
3737	CAC CAG AAG GAG CCC CCC TTC CTG TGG ATG GGC TAC GAG CTG CAC		
<hr/>			
	3260	3290	
3257	P D K W T V Q P I V L P E K D	NL4-3 genbank.seq	
3257	CCT GAT AAA TGG ACA GCA GTC CAG CCT ATA GTG CTG CCA GAA AAG GAC		
3255	P D K W T V Q P I V L P E K D	pNL4-3.seq	
3255	CCT GAT AAA TGG ACA GCA GTC CAG CCT ATA GTG CTG CCA GAA AAG GAC		
3782	P D K W T V Q P I V L P E K D	PHDMHgpm2.seq	
3782	CCC GAC AAG TGG ACC GTG CAG CCC ATC GTG CTG CCC GAG AAG GAC		
<hr/>			
	3320		
3302	S W T V N D I Q K L V G X L N	NL4-3 genbank.seq	
3302	AGC TGG ACT GTC ATT GAC ATA CAG AAA TTA GTG GGA AAA TTG AAT		
3300	S W T V N D I Q K L V G X L N	pNL4-3.seq	
3300	AGC TGG ACT GTC ATT GAC ATA CAG AAA TTA GTG GGA AAA TTG AAT		
3827	S W T V N D I Q K L V G X L N	PHDMHgpm2.seq	
3827	TCC TGG ACC GTG AAC GAC ATC CAG AAG CTG GTG GGC AAC CTG AAC		
<hr/>			
	3350	3380	
3347	W A S Q I Y A G I K V R Q L C	NL4-3 genbank.seq	
3347	TGG GCA AGT CAG ATT TAT GCA GGG ATT AAA GTC AGG CAA TTA TGT		
3345	W A S Q I Y A G I K V R Q L C	pNL4-3.seq	
3345	TGG GCA AGT CAG ATT TAT GCA GGG ATT AAA GTC AGG CAA TTA TGT		
3872	W A S Q I Y A G I K V R Q L C	PHDMHgpm2.seq	
3872	TGG GCC TCC CAG ATC TAC GCC GGC ATC AAA GTC CGC CAG CTG TGC		
<hr/>			
	3410		
3392	K L L R G T K A L T E V V P L	NL4-3 genbank.seq	
3392	AAA CTT CTT AGG GCA ACC AAA GCA CTA ACA GAA GTA GTC CCA CTA		
3390	K L L R G T K A L T E V V P L	pNL4-3.seq	
3390	AAA CTT CTT AGG GCA ACC AAA GCA CTA ACA GAA GTA GTC CCA CTA		
3917	K L L R G T K A L T E V V P L	PHDMHgpm2.seq	
3917	AAG CTG CTG CGC GGC ACC AAG GCC CTG ACC GAG GTG GTG CCC CTG		

Fig. 9E

18/29

Alignment Report of Codon Optimization (pol).MEG, using Clustal method with PAM250 residue weight table.

	3440														3470															
3437	T	E	E	A	E	L	E	L	A	E	N	R	E	I	L														NL4-3 genbank.SEQ	
3437	ACA	GAA	GAA	GCA	GAG	CTA	GAA	CTG	GCA	GAA	AAC	AGG	GAG	ATT	CTA														pNL4-3.seq	
3435	T	E	E	A	E	L	E	L	A	E	N	R	E	I	L															
3435	ACA	GAA	GAA	GCA	GAG	CTA	GAA	CTG	GCA	GAA	AAC	AGG	GAG	ATT	CTA															
3962	T	E	E	A	E	L	E	L	A	E	N	R	E	I	L														pHDMHgpm2.seq	
3962	ACC	GAG	GAG	GCC	GAG	CTG	GAG	CTG	GCC	GAG	AAC	CCC	GAG	ATC	CTG															
	3500																													
3482	K	E	P	V	H	G	V	Y	Y	D	P	S	K	D	L														NL4-3 genbank.SEQ	
3482	AAA	GAA	CGG	GTA	CAT	GGG	GTG	TAT	TAT	GAC	CCA	TCA	AAA	GAC	TTA															
3480	K	E	P	V	H	G	V	Y	Y	D	P	S	K	D	L														pNL4-3.seq	
3480	AAA	GAA	CGG	GTA	CAT	GGG	GTG	TAT	TAT	GAC	CCA	TCA	AAA	GAC	TTA															
4007	K	E	P	V	H	G	V	Y	Y	D	P	S	K	D	L														pHDMHgpm2.seq	
4007	AAG	GAG	CCC	GTG	CAC	GGC	GTG	TAC	TAC	GAC	CCC	TCC	AAG	GAC	CTG															
	3530														3560															
3527	I	A	E	I	Q	X	Q	G	Q	G	Q	W	T	Y	Q														NL4-3 genbank.SEQ	
3527	ATA	GCA	GAA	ATA	CAG	AAG	CAG	GGG	CRA	GCC	CRA	TGG	ACA	TAT	CRA															
3525	I	A	E	I	Q	K	Q	G	Q	G	Q	W	T	Y	Q														pNL4-3.seq	
3525	ATA	GCA	GAA	ATA	CAG	AAG	CAG	GGG	CRA	GCC	CRA	TGG	ACA	TAT	CRA															
4052	I	A	E	I	Q	K	Q	G	Q	G	Q	W	T	Y	Q														pHDMHgpm2.seq	
4052	ATC	GCC	GAG	ATC	CAG	RAG	CAG	GGC	CAG	GGC	CAG	TGG	ACC	TAC	CAG															
	3590																													
3572	I	Y	Q	E	P	F	K	N	L	K	T	G	K	Y	A														NL4-3 genbank.SEQ	
3572	ATT	TAT	CAA	GAG	CCA	TTT	AAA	AAU	CTG	AAA	ACA	GGA	AAA	TAT	GCA															
3570	I	Y	Q	E	P	F	K	N	L	K	T	G	K	Y	A														pNL4-3.seq	
3570	ATT	TAT	CAA	GAG	CCA	TTT	AAA	AAU	CTG	AAA	ACA	GGA	AAA	TAT	GCA															
4097	I	Y	Q	E	P	F	K	N	L	K	T	G	K	Y	A														pHDMHgpm2.seq	
4097	ATC	TAC	CAG	GAG	CCC	TTC	AAG	AAC	CTG	AAG	ACC	GGC	AAA	TAC	GCC															
	3620														3650															
3617	R	M	K	G	A	H	T	N	D	V	K	Q	L	T	E														NL4-3 genbank.SEQ	
3617	AGA	ATG	AAG	GGT	GCC	CAC	ACT	AAU	GAT	GTG	AAA	CAA	TTA	ACA	GAG															
3615	R	M	K	G	A	H	T	N	D	V	K	Q	L	T	E														pNL4-3.seq	
3615	AGA	ATG	AAG	GGT	GCC	CAC	ACT	AAU	GAT	GTG	AAA	CAA	TTA	ACA	GAG															
4142	R	M	K	G	A	H	T	N	D	V	K	Q	L	T	E														pHDMHgpm2.seq	
4142	CGC	ATG	AAG	GGC	GCC	CAC	ACC	AAC	GAC	GTG	AAG	CAG	CTG	ACC	GAG															
	3680																													
3662	A	V	Q	K	I	A	T	E	S	I	V	I	W	G	K														NL4-3 genbank.SEQ	
3662	GCA	GTA	CAA	AAA	ATA	GCC	ACA	GAA	AGC	ATA	GTA	ATA	TGG	GCA	AAG															
3660	A	V	Q	K	I	A	T	E	S	I	V	I	W	G	K														pNL4-3.seq	
3660	GCA	GTA	CAA	AAA	ATA	GCC	ACA	GAA	AGC	ATA	GTA	ATA	TGG	GCA	AAG															
4187	A	V	Q	K	I	A	T	E	S	I	V	I	W	G	K														pHDMHgpm2.seq	
4187	GCC	GTG	CAG	AAG	ATC	GCC	ACC	GAG	TCC	ATC	GTG	ATC	TGG	GCC	AAG															

Fig. 9F

19/29

Alignment Report of Codon Optimization (pol).MEG, using Clustal method with PAM250 residue weight table.

	3710	3740	
3707	T P K F K L P I Q K E T W E A		NL4-3 genbank.SEQ
3707	ACT CCT AAA TTT AAA TTA CCC ATA CAA AAG GAA ACA TGG GAA GCA		
3705	T P K F K L P I Q K E T W E A		pNL4-3.seq
3705	ACT CCT AAA TTT AAA TTA CCC ATA CAA AAG GAA ACA TGG GAA GCA		
4232	T P K F K L P I Q K E T W E A		
4232	ACT CCC AAG TTC AAG CTG CCC ATC CAG AAG GAG ACC TGG GAG GCC		pHDMHgpm2.seq
	3770		
3752	W W T E Y W Q A T W I P E W E		NL4-3 genbank.SEQ
3752	TGG TGG ACA GAG TAT TGG CAA GCC ACC TGG ATT CCT GAG TGG GAG		
3750	W W T E Y W Q A T W I P E W E		pNL4-3.seq
3750	TGG TGG ACA GAG TAT TGG CAA GCC ACC TGG ATT CCT GAG TGG GAG		
4277	W W T E Y W Q A T W I P E W E		
4277	TGG TGG ACC GAG TAC TGG CAG GCC ACC TGG ATC CCC GAG TGG GAG		pHDMHgpm2.seq
	3800	3830	
3797	F V N T P P L V K L H Y Q L E		NL4-3 genbank.SEQ
3797	TTT GTC AAT ACC CCT CCC TTA GTG AAG TTA TGG TAC CAG TTA GAG		
3795	F V N T P P L V K L H Y Q L E		pNL4-3.seq
3795	TTT GTC AAT ACC CCT CCC TTA GTG AAG TTA TGG TAC CAG TTA GAG		
4322	F V N T P P L V K L H Y Q L E		
4322	TTC GTG AAC ACC CCC CCC CTG GTG AAG CTG TGG TAC CAG CTG GAG		pHDMHgpm2.seq
	3860		
3842	K E P I I G A E T F Y V D S A		NL4-3 genbank.SEQ
3842	AAA GAA CCC ATA ATA GGA GCA GAA ACT TTC TAT GTA GAT GGG GCA		
3840	K E P I I G A E T F Y V D G A		pNL4-3.seq
3840	AAA GAA CCC ATA ATA GGA GCA GAA ACT TTC TAT GTA GAT GGG GCA		
4367	K E P I I G A E T F Y V D G A		
4367	AAG GAG CCC ATC ATC GGC GCC GAG ACC TTC TAC CTG GAC GGC GCC		pHDMHgpm2.seq
	3890	3920	
3887	A N R E T K L G K A G Y V T D		NL4-3 genbank.SEQ
3887	GCC AAT AGG GAA ACT AAA TTA GGA AAA GCA GGA TAT GTA ACT GAC		
3885	A N R E T K L G K A G Y V T D		pNL4-3.seq
3885	GCC AAT AGG GAA ACT AAA TTA GGA AAA GCA GGA TAT GTA ACT GAC		
4412	A N R E T K L G K A G Y V T D		
4412	GCC AAC CGC GAG ACC AAG CTG GCC AAG GCC GGC TAC GTG ACC GAC		pHDMHgpm2.seq
	3950		
3932	R G R Q X V V P L T D T T N Q		NL4-3 genbank.SEQ
3932	AGA GGA AGA CAA AAA GTT GTC CCC CTA ACN GAC ACA ACA AAT CAG		
3930	R G R Q X V V P L T D T T N Q		pNL4-3.seq
3930	AGA GGA AGA CAA AAA GTT GTC CCC CTA ACN GAC ACA ACA AAT CAG		
4457	R G R Q X V V P L T D T T N Q		
4457	CGC GGC CGC CAG AAG GTG GTG CCC CTG ACC GAC ACC AAC CAG		pHDMHgpm2.seq

Fig. 9G

20/29

Alignment Report of Codon Optimization (pol)MEG, using Clustal method with PAM250 residue weight table.

	3980	4010	
3977	K T E L Q A I H L A L Q D S G	NL4-3 genbank. SEQ	
3977	AAG ACT GAG TTA CAA GCA ATT CAT CTA GCT TTG CAG GAT TCG GGA		
3975	K T E L Q A I H L A L Q D S G	pNL4-3.seq	
3975	AAG ACT GAG TTA CAA GCA ATT CAT CTA GCT TTG CAG GAT TCG GGA		
4502	K T E L Q A I H L A L Q D S G	pHDMHgpm2.seq	
4502	AAG ACC GAG CTG CAG GCC ATC AAC CTG GCC CTG CAA GAC TCC GGC		
	4040		
4022	L E V N I V T D S Q Y A L G I	NL4-3 genbank. SEQ	
4022	TTA GAA GTA AAC ATA GTG ACA GAC TCA CAA TAT GCA TTG GGA ATC		
4020	L E V N I V T D S Q Y A L G I	pNL4-3.seq	
4020	TTA GAA GTA AAC ATA GTG ACA GAC TCA CAA TAT GCA TTG GGA ATC		
4547	L E V N I V T D S Q Y A L G I	pHDMHgpm2.seq	
4547	CTG GAG GTG AAC ATC GTG ACC GAC TCC CAG TAT GCA TTG GGC ATC		
	4070	4100	
4067	I Q A Q P D K S E S Z L V S Q	NL4-3 genbank. SEQ	
4067	ATT CAA GCA CAA CCA GAT AAG AGT GAA TCA GAG TTA GTC AGT CAA		
4065	I Q A Q P D K S E S E L V S Q	pNL4-3.seq	
4065	ATT CAA GCA CAA CCA GAT AAG AGT GAA TCA GAG TTA GTC AGT CAA		
4592	I Q A Q P D K S E S E L V S Q	pHDMHgpm2.seq	
4592	ATC CAG GCC CAG CCC GAC AAG TCC GAG TCC GAG CTG CTG TCC CAG		
	4130		
4112	I I E Q L I K K E K V Y L A W	NL4-3 genbank. SEQ	
4112	ATA ATA GAG CAG TTA ATA ATA AAA AAG GAA AAA GTC TAC CTG GCA TGG		
4110	I I E Q L I K K E K V Y L A W	pNL4-3.seq	
4110	ATA ATA GAG CAG TTA ATA ATA AAA AAG GAA AAA GTC TAC CTG GCA TGG		
4637	I I E Q L I K K E K V Y L A W	pHDMHgpm2.seq	
4637	ATC ATC GAG CAG CTG ATC AAG AAG GAG AAG GTG TAC CTG GCC TGG		
	4160	4190	
4157	V P A H R G I G G N E Q V D G	NL4-3 genbank. SEQ	
4157	GTA CCA GCA CAC AAA GGA ATT GGA GGA ATT GAA CAA GAA GAT GGG		
4155	V P A H R G I G G N E Q V D K	pNL4-3.seq	
4155	GTA CCA GCA CAC AAA GGA ATT GGA GGA ATT GAA CAA GAA GAT AAG		
4682	V P A H R G I G G N E Q V D K	pHDMHgpm2.seq	
4682	GTG CCC GCC CAC AAG GGC ATC GGC GGC AAC GAG CAG GTG GAC AAG		
	4220		
4202	L V S A G I R K V L F L D G I	NL4-3 genbank. SEQ	
4202	TTG GTC AGT GCT GGA ATC AGG AAA GAA CTA TTT TTA GAT GGA ATA		
4200	L V S A G I R K V L F L D G I	pNL4-3.seq	
4200	TTG GTC AGT GCT GGA ATC AGG AAA GAA CTA TTT TTA GAT GGA ATA		
4727	L V S A G I R K V L F L D G I	pHDMHgpm2.seq	
4727	CTG GTG TCC GCC GGC ATC CGC AAG GTG CTG TTC CTG GAC GGC ATC		

Fig. 9H

21/29

Alignment Report of Codon Optimization (pol).MEG, using Clustal method with PAM250 residue weight table.

	4250	4280	
4247	D K A Q E E H E K Y H S N W R	NL4-3 genbank.SEQ	
4247	GAT AAG GCC CAA GAA GAA CAT GAG AAA TAT CAC ACT AAT TGG AGA		
4245	D K A Q E E H E K Y H S N W R	pNL4-3.seq	
4245	GAT AAG GCC CAA GAA GAA CAT GAG AAA TAT CAC ACT AAT TGG AGA		
4772	D K A Q E E H E K Y H S N W R	pHDMHgpm2.seq	
4772	GAC AAG GCC CAG GAG GAG CAC GAG AAC TAC CAC TCC AAC AAC TGG CGC		
	4310		
4292	A M A S D F N L P P V V A K E	NL4-3 genbank.SEQ	
4292	GCA ATG CCT AGT GAT TTT AAC CTA CCA CCT GTC GTC GCA AAA GAA		
4290	A M A S D F N L P P V V A K E	pNL4-3.seq	
4290	GCA ATG CCT AGT GAT TTT AAC CTA CCA CCT GTC GTC GCA AAA GAA		
4817	A M A S D F N L P P V V A K E	pHDMHgpm2.seq	
4817	GCC ATG GCC TCC GAC TTC AAC CTG CCC CCC GTG GTG GCC AAG GAG		
	4340	4370	
4337	I V A S C D K C Q L K G E A M	NL4-3 genbank.SEQ	
4337	ATA GTA GCC AGC TGT GAT AAA TGT CAG CTA AAA GGG GAA GCC ATG		
4335	I V A S C D K C Q L K G E A M	pNL4-3.seq	
4335	ATA GTA GCC AGC TGT GAT AAA TGT CAG CTA AAA GGG GAA GCC ATG		
4862	I V A S C D K C Q L K G E A M	pHDMHgpm2.seq	
4862	ATC GTG GCC TCC TGC GAC AAG TGC CAG CTG AAG GCC GAG GCC ATG		
	4400		
4382	H G Q V D C S P G I W Q L D C	NL4-3 genbank.SEQ	
4382	CAT GGA CAA GTA GAC TGT AGC CCA GGA ATA TGG CAG CTA GAT TGT		
4380	H G Q V D C S P G I W Q L D C	pNL4-3.seq	
4380	CAT GGA CAA GTA GAC TGT AGC CCA GGA ATA TGG CAG CTA GAT TGT		
4907	H G Q V D C S P G I W Q L D C	pHDMHgpm2.seq	
4907	CAC GGC CAG GTG GAC TGC TCC CCC GGC ATC TGG CAG CTG GAC TGC		
	4430	4460	
4427	T H L E G K V I L V A V H V A	NL4-3 genbank.SEQ	
4427	ACA CAT TTA GAA GGA AAA GTT ATC TTG GTA GCA GTT CAT GTA GCC		
4425	T H L E G K V I L V A V H V A	pNL4-3.seq	
4425	ACA CAT TTA GAA GGA AAA GTT ATC TTG GTA GCA GTT CAT GTA GCC		
4952	T H L E G K V I L V A V H V A	pHDMHgpm2.seq	
4952	ACC CAC CTG GAG GCC AAG GTG ATC CTG GTG GCC GTG CAC GTG GCC		
	4490		
4472	S G Y I E A E V I P A E T G Q	NL4-3 genbank.SEQ	
4472	AGT GGA TAT ATA GAA GCA GAA GAA ATT CCA GCA GAG ACA GGG CAA		
4470	S G Y I E A E V I P A E T G Q	pNL4-3.seq	
4470	AGT GGA TAT ATA GAA GCA GAA GAA ATT CCA GCA GAG ACA GGG CAA		
4997	S G Y I E A E V I P A E T G Q	pHDMHgpm2.seq	
4997	TCC GCC TAC ATC GAG GCC GAG GTG ATC CCC GCC GAG ACC GCC CAG		

Fig. 9I

22/29

Alignment Report of Codon Optimization (pol) MEG, using Clustal method with PAM250 residue weight table.

	4520	4550	
4517	E T A Y F L L K L A G R W P V	NL4-3 genbank.seq	
4517	GAA ACA GCA TAC TTC CTC TTA AAA TTA GCA GGA AGA TGG CCA GTA		
4515	E T A Y F L L K L A G R W P V	pNL4-3.seq	
4515	GAA ACA GCA TAC TTC CTC TTA AAA TTA GCA GGA AGA TGG CCA GTA		
5042	E T A Y F L L K L A G R W P V	pHDHgpm2.seq	
5042	GAG ACC GCC TAC TTC CTG CTG AAG CTG GCC GGC CGC TGG CCC GTG		
<hr/>			
	4580		
4562	K T V H T D N G S M E T S T T	NL4-3 genbank.seq	
4562	AAA ACA GTA CAT ACA GAC AAT GCC AGC AAT TTC ACC AGT ACT ACA		
4560	K T V H T D N G S M E T S T T	pNL4-3.seq	
4560	AAA ACA GTA CAT ACA GAC AAT GCC AGC AAT TTC ACC AGT ACT ACA		
5087	K T V H T D N G S M E T S T T	pHDHgpm2.seq	
5087	AAG ACC GTG CAC ACC GAC AAC GCC TCC AAC TTC ACC TCC ACC ACC		
<hr/>			
	4610	4640	
4607	V K A A C W W A G I K Q E F G	NL4-3 genbank.seq	
4607	GTT AAG GCC GCC TGT TGG TGG GCG GGG ATC AAG CAG CAG GAA TTT CCC		
4605	V K A A C W W A G I K Q E F G	pNL4-3.seq	
4605	GTT AAG GCC GCC TGT TGG TGG GCG GGG ATC AAG CAG CAG GAA TTT CCC		
5132	V K A A C W W A G I K Q E F G	pHDHgpm2.seq	
5132	GTC AAG GCC GCC TGC TGG TGG GCC GGC ATC AAG CAG CAG GAG TTC GGC		
<hr/>			
	4670	4730	
4652	I P Y N P Q S Q G V I E S M N	NL4-3 genbank.seq	
4652	ATT CCC TAC AAT CCC CAA AGT CAA GGA GTA ATA GAA TCT ATG AAT		
4650	I P Y N P Q S Q G V I E S M N	pNL4-3.seq	
4650	ATT CCC TAC AAT CCC CAA AGT CAA GGA GTA ATA GAA TCT ATG AAT		
5177	I P Y N P Q S Q G V I E S M N	pHDHgpm2.seq	
5177	ATC CCC TAC AAC CCC CAG TCC CAG GGC GTG ATC AAG CAG TCC ATG AAC		
<hr/>			
	4760		
4697	R E L K K I I G Q V R D Q A E	NL4-3 genbank.seq	
4697	AAA GAA TTA AAG AAA ATT ATA GGA CAG GTA AGA GAT CAG GCT GAA		
4695	R E L K K I I G Q V R D Q A E	pNL4-3.seq	
4695	AAA GAA TTA AAG AAA ATT ATA GGA CAG GTA AGA GAT CAG GCT GAA		
5222	R E L K K I I G Q V R D Q A E	pHDHgpm2.seq	
5222	AAG GAG CTG AAG AAG ATC ATC GGC CAA GTC CGC GAC CAG GCC GAG		
<hr/>			
	4760		
4742	H L K T A V Q M A V F I H N F	NL4-3 genbank.seq	
4742	CAT CTT AAG ACA GCA GTA CAA ATG GCA GTA TTC ATC CAC AAT TTT		
4740	H L K T A V Q M A V F I H N F	pNL4-3.seq	
4740	CAT CTT AAG ACA GCA GTA CAA ATG GCA GTA TTC ATC CAC AAT TTT		
5267	H L K T A V Q M A V F I H N F	pHDHgpm2.seq	
5267	CAC CTG AAG ACC GCC GTG CAG ATG GCC GTG TTC ATC CAC AAC TTC		

Fig. 9J

23/29

Alignment Report of Codon Optimization (pol) MEG, using Clustal method with PAM250 residue weight table.

	4790	4820	
4787	K R K G G I G G Y S A G E R I	NL4-3 genbank.seq	
4787	AAA AGA AAA GGG GGG ATT GGG GGG TAC AGT GCA GGG GAA AGA ATA		
4785	K R K G G I G G Y S A G E R I	pNL4-3.seq	
4785	AAA AGA AAA GGG GGG ATT GGG GGG TAC AGT GCA GGG GAA AGA ATA		
5312	K R K G G I G G Y S A G E R I		
5312	AAG CGC AAG GCC GGC ATC GCC GGC TAC TCC GCC GGC GAG CGC ATC	pHDMHgpm2.seq	
	4850		
4832	V D I I A T D I Q T K E L Q R	NL4-3 genbank.seq	
4832	GTA GAC ATA ATA GCA ACA GAC ATA CAA ACT AAA GAA TTA CAA AAA		
4830	V D I I A T D I Q T K E L Q R	pNL4-3.seq	
4830	GTA GAC ATA ATA GCA ACA GAC ATA CAA ACT AAA GAA TTA CAA AAA		
5357	V D I I A T D I Q T K E L Q R		
5357	GTG GAC ATC ATC GCC ACC GAC ATC CAG ACC RAG GAG CTG CAG AAG	pHDMHgpm2.seq	
	4880	4910	
4877	Q I T K I Q N F R V Y Y R D S	NL4-3 genbank.seq	
4877	CAA ATT ACA AAA ATT CAA ATT TTT CGG GTT TAT TAC AGG GAC AGC		
4875	Q I T K I Q N F R V Y Y R D S	pNL4-3.seq	
4875	CAA ATT ACA AAA ATT CAA ATT TTT CGG GTT TAT TAC AGG GAC AGC		
5402	Q I T K I Q N F R V Y Y R D S		
5402	CAG ATC ACC AAG ATC CAG AAC TTC CGC GTG TAC TAC CGC GAC TCC	pHDMHgpm2.seq	
	4940		
4922	R D P V W K G P A K L L W K G	NL4-3 genbank.seq	
4922	AGA GAT CCA GTT TGG AAA GGA CCA GCA AAG CTC CTC TGG AAA GGT		
4920	R D P V W K G P A K L L W K G	pNL4-3.seq	
4920	AGA GAT CCA GTT TGG AAA GGA CCA GCA AAG CTC CTC TGG AAA GGT		
5447	R D P V W K G P A K L L W K G		
5447	CGC GAC CCC GTG TGG AAG GGC CCC GCC AAG CTG CTG TGG AAG GGC	pHDMHgpm2.seq	
	4970	5000	
4967	E G A V V I Q D N S D I R V V	NL4-3 genbank.seq	
4967	GAA GGG GCA GTA GTA ATA CAA GAT ATT AGT GAC ATA AAA GTA GTG		
4965	E G A V V I Q D N S D I K V V	pNL4-3.seq	
4965	GAA GGG GCA GTA GTA ATA CAA GAT ATT AGT GAC ATA AAA GTA GTG		
5492	E G A V V I Q D N S D I R V V		
5492	GAG GGC GCC GTG GTG ATC CAG GAC AAC TCC GAC ATC AAG GTG GTG	pHDMHgpm2.seq	
	5030		
5012	P R R K A K I I R D Y G K Q M	NL4-3 genbank.seq	
5012	CCA AGA AGA AAA GCA AAG ATC ATC AGG GAT TAT GGA AAA CAG ATG		
5010	P R R K A K I I R D Y G K Q M	pNL4-3.seq	
5010	CCA AGA AGA AAA GCA AAG ATC ATC AGG GAT TAT GGA AAA CAG ATG		
5537	P R R K A K I I R D Y G K Q M		
5537	CCC CGC CGC AAG GCC AAG ATC ATC CGC GAC TAC GGC AAG CAG ATG	pHDMHgpm2.seq	

Fig. 9K

26/29

Alignment Report of Codon Optimization (pol).MEG, using Clustal method with PAM250 residue weight table.

	5060	5090	
5057	A G D D C V A S R Q D E D		NL4-3_genbank.SEQ
5057	GCA GGT GAT GAT TGT GTG GCA AGT AGA CAG CAT GAG GAT TAA		
5055	A G D D C V A S R Q D E D		pNL4-3.seq
5055	GCA GGT GAT GAT TGT GTG GCA AGT AGA CAG CAT GAG GAT TAA		
5582	A G D D C V A S R Q D E D		pHDMHgpm2.seq
5582	GCC GGC GAC GAC TGC GTG GCC TCC CGC CAG GAC GAG GAC TAA		

Fig. 9L

AGCTTGGCCC	ATTGCATACG	TTGTATCCAT	ATCATAATAAT	GTACATTAT	ATTGGCTCAT	60
GTCCAACATT	ACCGCCATGT	TGACATTGAT	TATTGACTAG	TTATTAATAG	TAATCAATTA	120
CGGGGTCATT	AGTTCATAGC	CCATATATGG	AGTTCCCGT	TACATAACTT	ACGGTAAATG	180
GCCCCCTGG	CTGACCGGCC	AACGACCCC	GCCCATTGAC	GTCAATAATG	ACGTATGTTG	240
CCATAGAAC	GCCAATAGGG	ACTTTCCATT	GACGTCAATG	GGTGGAGTAT	TTACGGTAAA	300
CTGCCCACTT	GGCAGTACAT	CAAGTGATC	ATATGCCAAG	TACGGCCCT	ATTGACGTCA	360
ATGACGGTAA	ATGGCCCGCC	TGGCATTATG	CCCAAGTACAT	GACCTTATGG	GACTTCCCTA	420
CTTGGCAGTA	CATCTACGTA	TTAGTCATG	CTATTACCAT	GGTGTGCGG	TTTGGCAGT	480
ACATCAATGG	GCGTGGATAG	CGGTTTGACT	CACGGGATT	TCCAAGTCTC	CACCCCATTG	540
ACGTCATGG	GAGTTGTTT	TGGCACCAA	ATCAACGGGA	CTTTCACAAA	TGTGCTAACAA	600
ACTCCGCC	ATTGACGCAA	ATGGCGGT	GGCGTGTACG	GTGGGAGGTC	TATATAAGCA	660
GAGCTCGIT	AGTGAACCGT	CAGATGACAT	GGAGACGCCA	TCCAGCTGT	TTTGACCTCC	720
ATAGAAGACA	CCCGGACCGA	TCCAGCTCC	CCTCGAAGCT	GATCCTGAGA	ACTTCAGGGT	780
GAGTCTATGG	GACCTTGTATG	GTTCCTTTC	CCCTTCTTTT	CTATGGTTAA	GTTCATGTCA	840
TAGGAAGGGG	AGAAGTAACA	GGGTACACAT	ATTGACCAA	TCAGGGTAAT	TTTGATTTG	900
TAATTTTAA	AAATGCTTTC	TTCTTTTAA	ATACTTTT	GTTCATCTTA	TTTCATAATAC	960
TTTCCCTAA	CTCTTTCTT	CAGGGCAATA	ATGATACAT	GTATGATGCC	TCTTGGCAC	1020
ATTC TAAAGA	ATAACAGTGA	TAATTTCTGG	GTAAAGCCAT	TAGCAATATT	TCTGCATATA	1080
AATTTCTG	CATATAAAAT	GTAAACTGATG	TAAGAGGTTT	CATATTGCTA	ATASCAGCTA	1140
CAATCCAGCT	ACCATTCTGC	TTTTATTTA	TGGTTGGGAT	AAGGCTGGAT	TATTCCTGAGT	1200
CCAAAGCTAGG	CCCTTTTCT	AATCATGTC	ATACCTCTTA	TCTTCCTCCC	ACAGCTCTG	1260
GGCACACGTG	TGGTCTGTG	GCTGGCCAT	CACTTGGCA	AAGAATTCTA	GACTGCCATG	1320
GGCCCCCGG	CCTCCGTCT	GTCCGGCGC	CAGCTGGACA	AGTGGGAGAA	GATCCGCCTG	1380
CGCCCCCGG	GCAAGAACCA	GTACAAGCTG	AAGCACATCG	TGTGGCCCT	CCCGGAGCTG	1440
GAGCCCTT	CGCTGAACCC	CGGCCCTCTG	GAGACCTCCG	AGGGCTGCCG	CCAGATCTG	1500
GGCCAGCTGC	ACCCCTCCCT	SCAACACGGC	TCCGAGGAGC	TGGCCTCCCT	GTACAACACC	1560
ATGCCCGTGC	TGTACTGGT	GCACCAAGGC	ATCGACGTGA	AGGACACCAA	GGAGGCCCTG	1620
GACAGAGATG	AGGAGGAGCA	GAACAACTG	AAGAAGAAGG	CCAGCAGCC	CGCCGCCGAC	1680
ACCGGCAACA	ACTCCCAGGT	GTACCAAGAC	TACCCCATCG	TGCGAAACCT	GCAGGGCCAG	1740
ATGGTGCACC	AGGCCATCTC	CCCCCGCACC	CTGAACGCC	GGGTGAAGGT	GGTGGAGGAG	1800
AAGGCCCTCT	CCCCCGRAAGT	CATCCCCATG	TTCTCCGCCC	TGTCGGAGGG	CGCCACCCCC	1860
CAGGACCTGA	ACACCATGCT	GAACACCGT	GGCGGCCACCC	AGGAGCCCAT	GCAGATGCTG	1920
AAGGAGACCA	TCAACCGAGG	GGCCGCGAG	TGGGACCGC	TGCAACCCCGT	GCACCCCGC	1980
CCCATCGCCC	CCGGCCAGAT	GGCGGAGCC	CGCGCTCCG	ACATCGCCCG	CACCCACCTCC	2040
ACCTTGCAGG	ACCAAGATCGG	CTGGATGACC	CACACCCCCC	CCATCCCCGT	GGGGAGATC	2100
TACAAGCGCT	GGATCATCTC	GGCCCTGAAC	AAGATCGTG	GCATGACTC	CCCCACCTCC	2160
ATCCGTGACA	TCGGCCAGGG	CCCAAAGGAG	CCCTCCCGC	ACTACGTGGA	CCGCTTCTAC	2220
AAGACCCCTGC	GGCCGGAGCA	GGCCTCCAG	GAGGTAAAGA	ACTGGATGAC	CGAGACCTG	2280
CTGGTGCAGA	ACGCCAACCC	CGACTGCAAG	ACCATCCTGA	AGGCCCTGGG	CCCCGGCC	2340
ACCCCTGGAGG	AGATGATGAC	CGCCCTCCAG	GGCGTGGGGC	GGCCCGGCCA	CAAGGGCC	2400
GTGCTGGCCG	AGGCCATGTC	CCAAGTCACC	AACCCGCCA	CCATCATGAT	CCAGAAGGGC	2460
AACTTCCGCA	ACCAAGCGAA	GACCGTGAAG	TGCTTCAACT	GGCGCAAGGA	GGGCCACATC	2520
GCCAAAGAATC	GGCGCGCCCC	CGCGAAGAAG	GGCTGCTGGA	AGTGCAGGCAA	GGAGGGCCAC	2580
CAGATGAAAG	ATTTGACTGA	GAGACAGGCT	AAATTTTATG	GGAGAGATCTG	GCCTTCCCCAC	2640
AACGGAAAGGC	CAGGGAAATT	TGTCAGAGC	AGACCAAGAGC	CAACAGCCCC	ACCAGAAGAG	2700
AGCTTCAGGT	TTGGGAAAGA	GACAACAACT	CCCTCTCAGA	AGCAGGACCC	GATAGACAG	2760
GAACGTGATC	CTTAGCTTC	CCTCAGATCA	CTCTTGGCA	GGCACCCCTC	GTACACARTAA	2820

Fig. 10A

26/29

AGATCGGTGG	CCAGCTGAAG	GAGGCCCTGC	TGGACACCGG	CGCCGACGAC	ACCGTGCTGG	2880
AGGAGATGAA	CCTGCCCGC	CGCTGGAAGC	CCAAGATGAT	CGGCGCATC	GGCGGCTTCA	2940
TCAAAAGTCGG	CCGACTACGAC	CAGATCTCGA	TCGAGATCAG	CGGCCAACAG	GCCATCGGCA	3000
CCGTGCTGGT	GGGCCCCACC	CCCGTGAACA	TCATCGGCG	CAACCTGCTG	ACCCAGATCG	3060
GCTCACCCCT	GAACCTCCCC	ATCTCCCCA	TCGAGACCGT	GCCCGTGAAG	CTGAAGCCCC	3120
GCATGGACGG	CCCCAAAGTC	AAGCAGTGGC	CCCTGACCGA	GGAGAAGATC	AAGGCCCTGG	3180
TGGAGATCTG	CACCGAGATG	GAGAAGGAGG	CCAAGATCTC	CAAGATCGGC	CCCGAGAACCC	3240
CCTACAACAC	CCCCCTGTPC	GCCATCAAC	AGAAGGACTC	CACCAAGTGG	CGCAAGGCTGG	3300
TGGAACTCCG	CGAGCTGAAC	AAGCGCACCC	AGGACTCTCG	GGAGGTGCAAG	CTGGGCATCC	3360
CCCACCCCGC	CGGCGTGAAG	CAGAAGAAGT	CCGTGACCGT	GCTGGACGTG	GGCGACGCC	3420
ACTTCTCGT	GGCCCTGGAC	AAGGACTTCC	CCAAGTACAC	CGCCCTTCACC	ATCCCCCTCCA	3480
TCAACAACGA	GACCCCCGGC	ATCCGCTACC	AGTACAAACGT	GCTGCTTCAG	GGCTTGGAAAGG	3540
GCTCCCCCGC	CATCTCCAG	TGCTCCATGA	CCAAGATCCT	GGACCCCTTC	CGCAAGGAGA	3600
ACCCCGACAT	CGTGTATCTAC	CACTACATGG	ACGACCTGTA	CGTGGGCTCC	GACCTGGAGA	3660
TCGCCCCAGA	CCCCACCAAG	ATCGAGGAGC	TGCGCCAGCA	CCTGCTGCGC	TGGGGCTTCA	3720
CCACCCCCCGA	CAAGAAGGAGC	CCCCCTTCCCT	GTGGATGGGC	TACGAGCTGC	3780	
ACCCCGACAA	GTGGACCGTG	CAGCCCATCG	TGCTGCCCGA	GAAGGACTCC	TGACCGGTGA	3840
ACGACATCCA	GAAGCTGGTG	GGCAAGCTGA	ACTGGGCTC	CCAGATCTAC	GGCGGCATCA	3900
AAAGTCGCCA	GCTGTGCAAG	CTGCTGCCG	CCACCAAGGC	CCTGACCCAG	GTGGTGGCCCC	3960
TGACCGAGGA	GGCCCGACCTG	GAGCTGCCG	AGAACCGCGA	GATCTGAAG	GASCCCCGTGC	4020
ACGGGTGTA	CTACGACCCC	TCCAAGGACC	TGATGCCGGA	GAATCCAGAG	CAGGGCCAGG	4080
GCCAGTGGAC	CTACCAAGTC	TACCAAGGAGC	CTTCAAGAA	CTGTGAAGACC	GGCAAATACG	4140
CCCCCATGAA	GGCCGCCAAC	ACCAACCGAC	TGAAGCAGCT	GACCGAGGCC	GTGCCAGAAGA	4200
TGCCCCACCGA	GTCCATCTG	ATCTGGGG	AGACTCCCA	GTTCAGGCTG	CCCATCCAGA	4260
AGGAGACCTG	GGAGGCGCTGG	TGAGGACCTG	ACTGGCAGGC	CACCTGGATC	CCCGAGTGGG	4320
AGTTCTGTAA	CACCCCCCCC	CTGGTGAAGC	TGTGGTACCA	GCTGGAGAAG	GAGCCCATCA	4380
TCGCGGCCGA	GACCTCTAC	GTGGACGGCG	CGGCCAACCG	CGAGACCAAG	CTGGGCAAGG	4440
CCGGCTACGT	GACCGACCGC	GGCGCCAGA	AGTGCTGCC	CTTGACCCGAC	ACCAACCAAC	4500
AGAACACCGA	GCTGCAGGCC	ATGCCACCTG	CCCTCCAAGA	CTCGGGCTG	GACGTGAACA	4560
TCGTGACCGA	CTCCCAGTAT	GCATTGGCA	TCATCCAGGC	CCAGCCGAC	AAGTCCGAGT	4620
CCGAGCTGGT	GTCCCAGATC	ATCGAGCAGC	TGATCAAGAA	GGAGAAGGTG	TACCTGGCCT	4680
GGGTGCCCGC	CCACAAGGGC	ATCGGACGGC	AGGAGCAGGT	GGACAAGCTG	GTGTCGGCCG	4740
GCATCGCAA	GGTGTGTT	CTGGACGCC	TCGACAAGGC	CCAGAGGAG	CACCGAGAAGT	4800
ACCACTCAA	CTGGCGGCC	ATGGCCCTCG	ACTTCAACCT	GCCCCCGTG	GTGGCCAGG	4860
AGATGTGCC	CTCTCTGCCAC	ATGTGCCAGC	TGAAGGGCGA	GGCCATGAC	GGCCAGGTGG	4920
ACTGCTGCC	CGGCATCTGG	CAGCTGGACT	SCACCCACCT	GGAGGCGAAC	GTGATCTGG	4980
TGGCGGTGCA	CTGGCCCTCC	GGCTACATCG	AGGCGCAGGT	GATCCCCGCC	GAGACCGGCC	5040
AGGAGACCGC	CTACTCTCG	CTGAAGCTGG	CGGGCGCTG	CCCCGTGAAG	ACCGTGCACA	5100
CCGACAACCG	CTCCAACCTC	ACCTCAACCA	CCGTGAAGGC	CGGCTGCTGG	TGGCCGGCCA	5160
TCAAGCAGGA	GTTCGGCATC	CCCTACAAAC	CCAGTCCCA	GGCGGTGATC	GAGTCCATGA	5220
ACAAGGAGCT	GAAGAAGATC	ATGGGCCAG	TCCGGCAGCA	GGCGGAGCAC	CTGAAGACCG	5280
CCGIGCAGAT	GGCCGTGTPC	ATCCACAACT	TCAAGCGAA	GGGGGGCATC	GGGGCTACT	5340
CCGGCGGCCG	GCCCATCTG	GACATCATCG	CCACCGACAC	CCAGACCAAG	GACCTGAGA	5400
AGCAAGATCAC	CAAGATCCAG	ACTTCCCGT	TGTACTACG	CGACTCCCCG	GACCCCTGTG	5460
GGAAAGGCC	GGCCAAAGCTG	CTGTGGAGG	GGAGGGCGC	CGTGGTGAATC	CACCGACAAT	5520
CCGACATCAA	GGTGTGCCCC	GGCCCGAAGG	CCAAGATCAT	CCGGCACTAC	GCCAAGCAGA	5580
TGGCGGCCGA	CGACTGCGTG	GCCTCCCCCC	AGGACGAGGA	CTAACACATG	AAAAAGATTA	5640

Fig. 10B

27/29

GTAAAACACC ATAGGCCGCT CTAGAGGATC CAAGCTTATC GATAACCGTCG ACCTCGAGGG 5700
 CCCAGATCTA ATTACCCCCA CCAGTCAGG CTGCTTATCA GAAGTCGGT GCTGGTGTGG 5760
 CTAATGCCCT GGCCACAAG TATCACTAG CTGCTTTCT TCTGTCCAA TTTCATTAA 5820
 AGGTCTTCTT GTTCCCTAAG TCCAACACT AAACCTGGGG ATATTAATGAA GGGCCCTTGAG 5880
 CATCTGATT CTGCTTAATA AAAAACATT ATTTCATTG CAATGATGTA TTAAATTAT 5940
 TTCTGAATAT TTACTAAAAA AGGGATGTG GGAGGTCAGT GCATTTAAAAA CATAAAGAAA 6000
 TGAAGAGCTA GTTCAACCT TGAAAATAA CACTATATCT TAATCTCAT GAAAGAAGGT 6060
 GAGGTGCAAA ACAGCTAATG CACATGGCA ACAGCCCCCTG ATGCTCTATGC CTTATTCATC 6120
 CCTCAGAAA GGATTCAGT AGAGGCTTGA TTGGAGGTT AAAGTTTGC TATGCTGTAT 6180
 TTTCACATTAC TTATTTCTT AGCTGTCCTC ATGAATGTC TTTCACTACC CATTGCTTA 6240
 TCTCGCATCT CTGACCCCTG GTTCTCTTGC TTAGACATAC CACCTTCCC 6300
 CTGAAGCTGT CTTCCATGT TTACGGCGA GATGTTCTC CTCGCTTGG CCACCTAGCC 6360
 TTAGTTGCTC CTGTTGCTT ATAGAGCTC ACTTGAAGAA GGGAAAACAG GGGGCATGGT 6420
 TTGACTGTCC TGTGASCCCT TCTTCCCTGC CTCCCCCACT CACAGTGCAC CGGAATCCCT 6480
 CGACATGCCA GTCTAGATCA TTCTTGAAGA CGAAAGGGCC TCCTGTATAC CCTTATTTTA 6540
 TAGTTAATG TCATGATAAT AATGGTTCTC TAGACGTCAAG GTGGCATTT TCGGGAAAT 6600
 GTGCGCGGAA CCCCTATTTG TTATTTTC TAAATACATT CAATATGTA TCCGCTCATG 6660
 AGACAATAAC CCTGATAAAAT GCTCAATAA TATTGAAAAA GGAAAGACTAT GAGTATTCAA 6720
 CATTCCCTG TGCCCTTAT GCTCTTCTTGC GCGGCATTTT GCGCTCTGTG TTTTGCCTCAC 6780
 CCAGAACGCG TGGTAAAGT AAAAGATGCT GAAGATCACT TGCGTCCACG ACTGGGTTAC 6840
 ATCGAACTGG ATCTAACAG CGGTAAAGATC CTTGAGAGTT TTGCCCCGA AGAACGTTT 6900
 CCAATGATGA GCACTTTAA AGTTCTGCTA TGTCGGCGGG TATTATCCCG TATTGACGCC 6960
 GGGCAAGAGG AACTCGTCG CGCGATACAC TATCTCTCAGA ATGACTCTGT TGAGTACTCA 7020
 CCAGTCACAG AAAAGCATCT TACGGATGGC ATGACAGTAA GAAATTATG CAGTCTGCC 7080
 ATAACCATGA GTGATAACAC TGCGGCCAAC TTACTCTGA CAACGATCGG AGGACCGAAG 7140
 GAGCTAACCG CTTTTTGCA CAACATGGGG GATCATGTA CTGCTTGA TGTTGGAA 7200
 CGGGAGCTGA ATGAAAGCCAT ACCAAACGAG GAGCGTGCACCAACGATGCC TGAGCAATG 7260
 GCAACACTG TGCGAAACT ATTAACCTGC GAACACTTA CTCTAGCTC CGGGCAACAA 7320
 TTAAATAGCT CGATGGAGGG CGATAAAGTT GCAGGACAC TTCTGCTCTC GGGCTTCCG 7380
 GCTGGCTGGT TTATTCGTGA TAAATCTGGA GCGGTGAGC GTGGCTCTCG CGGTATCATT 7440
 GCAGCACTGG GCCCAGATGG TAAGCCTCTC CGTATCTGTAG TTATCTACAC GACGGGGAGT 7500
 CAGGCAACTA TGGATGAAACG AAATAGACAG ATCCGTGAGA TAGGTGCTC ACTGATTAAG 7560
 CATTGGTAAC TGTCAGACCA AGTTACTCA TATATACATT AGATTGATTT AAAACTTCAT 7620
 TTAAATTTA AAAGGATCTA GTGAGATCCTA TTGTTGATA ATCTCATGAC CAAAATCCCT 7680
 TAAAGTGTAGT TTCTGTTCCG CTGAGCTGCA GACCCCGTAG AAAAGATCAA AGGATCTTCT 7740
 TGAGATCCCTT TTCTCTGCG CGTAACTCTGC TGCTTGCAAA AAAAACC ACCGCTACCA 7800
 GCGGTGGTTT GTTGGCCGGA TCAAGAGCTA CCAACTCTT TTGGAAGGT AACCTGCTC 7860
 AGCAGAGCGC AGATACCAAA TACTGTTCTT CTAGTGTACG CGTAGTTAGG CCACCACTTC 7920
 AAGAACTCTG TAGCACCGCC TACATACCTC GCTCTGCTAA TCTCTGCTAC AGTGGCTGCT 7980
 GCGAGGGCG ATAAGTCGTG TCTTACCGGG TTGGACTCAA GACGATAGT ACCGGATAAG 8040
 GCGCAGCGGT CGGGCTGAAAC GGGGGGTTCG TGCAACACG CCGCTTGGA CGAACGACC 8100
 TACACCGAAC TGAGATACCT ACAGCGTGAG CTATGAGAAA GCGCCACGCT TCCCGAAGGG 8160
 AGAAAGCGGG ACAGGTATCC GTAAACCGGG AGGGTCGGAA CAGGAGACG CACGAGGGAG 8220
 CTCCAGGGG GAAACGCCCTG GTATCTTAT AGTCTCTGC GGTTTGCCTA CCTCTGACTT 8280
 GAGCTGCGAT TTCTGATG CTGCTCAGGG GGGCGGAGCC TATGGAAAAA CGCCACCAAC 8340
 GGATGCGCCG CGTGCCTG CTGGAGATGG CGGACCGCAT GGATATGTC TGCCAAAGGGT 8400
 TGTTTGCCTG ATTACACAGTT CTCCGCAAGA ATTGATTGGC TCCAATTCTT GGAGTGGTGA 8460

Fig. 10C

28/29

ATCCGTTAGC GAGGTGCCGC CGGCTTCAT TCAGGTGAG GTGGCCCGGC TCCATGCACC 8520
GGGACGCAAC GCGGGGAGCC AGACAAGTA TAGGGGGCG CCTACAATCC ATGCCAACCC 8580
GTTCATCTG CTCGCCGAGG CGGCATAAAT CCCCGTGACG ATCAGCGGT CAATGATCGA 8640
AGTTAGGCTG GTAAGAGCGG CGAGCGATCC TTGAACCTGT CCCTGATGGT CGTCATCTAC 8700
CTGCCCTGGAC ACCATGGCCT GCAACCGGGG CATCCCAGTC CGGCCGGAAAG CGAGAAGAAAT 8760
CATATAATGGGG AAGGCCATCC ACCCTCGGT CGGGGAGCTT TTGCAAAG CCTAGGCTC 8820
CAAAAAAGCC TCCTCACTAC TTCTGGAATA CCTCAGAGGC CGAGGCGGCC TCGGCCCTCG 8880
CATAAAATAAA AAAAATTAGT CAGCCATG 8908

Fig. 10D

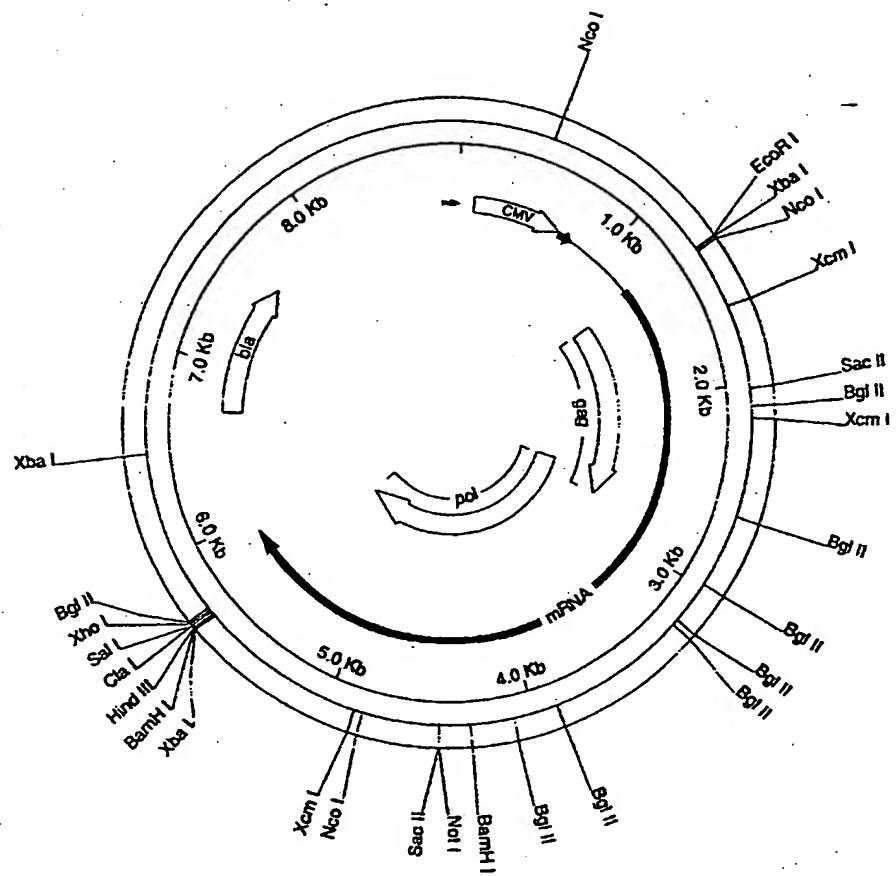


Fig. 11

INTERNATIONAL SEARCH REPORT

1	National Application No PCT/US 99/20675
---	---

A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/86 C12N5/10 C12N7/04 C12N15/49 C07K14/16				
According to International Patent Classification (IPC) or to both national classification and IPC				
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 C12N C07K				
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched				
Electronic data base consulted during the international search (name of data base and, where practical, search terms used)				
C. DOCUMENTS CONSIDERED TO BE RELEVANT				
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.		
A	<p>NALDINI L ET AL: "IN VIVO GENE DELIVERY AND STABLE TRANSDUCTION OF NONDIVIDING CELLS BY A LENTIVIRAL VECTOR" SCIENCE, US, AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE., vol. 272, no. 5259, 12 April 1996 (1996-04-12), pages 263-267, XP000583652 ISSN: 0036-8075 cited in the application the whole document</p> <p align="center">-/-</p>	1-4		
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C.		<input type="checkbox"/> Patent family members are listed in annex.		
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the International filing date "L" document which may throw doubts on priority, claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the International filing date but later than the priority date claimed				
Date of the actual completion of the international search 25 February 2000		Date of mailing of the international search report 03/03/2000		
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patenttaan 2 NL - 2200 HV Rijswijk Tel: (+31-70) 340-2040, Fax: 31 851 890 41, Fac: (+31-70) 340-3016		Authorized officer Chambonnet, F		

INTERNATIONAL SEARCH REPORT

National Application No
PCT/US 99/20675

C(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT	
Category	Citation of document, with indication where appropriate, of the relevant passages Relevant to claim No.
A	<p>HASELHORST D ET AL: "STABLE PACKAGING CELL LINES AND HIV-1 BASED RETROVIRAL VECTOR SYSTEMS" GENE THERAPY, GB, MACMILLAN PRESS LTD., BASINGSTOKE, vol. 1, no. SUPPL. 02, 18 November 1994 (1994-11-18), page S14 XP002063698 ISSN: 0969-7128 the whole document</p>
A	<p>ST LOUIS D ET AL: "CONSTRUCTION AND CHARACTERIZATION OF HIV-1 RETROVIRAL VECTORS AND REPPLICATION-DEFECTIVE HIV-1 PACKAGING CELL LINES" INTERNATIONAL CONFERENCE ON AIDS AND THE STD WORLD CONGRESS, XX, XX, 1 June 1993 (1993-06-01), page 244 XP002063695 the whole document</p>
A	<p>CARROLL R ET AL: "A HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1)-BASED RETROVIRAL VECTOR SYSTEM UTILIZING STABLE HIV-1 PACKAGING CELL LINES" JOURNAL OF VIROLOGY, US, THE AMERICAN SOCIETY FOR MICROBIOLOGY, vol. 68, no. 9, 1 September 1994 (1994-09-01), pages 6047-6051, XP002063697 ISSN: 0022-538X the whole document</p>
X	<p>HOLLER T P ET AL: "HIV1 INTEGRASE EXPRESSED IN ESCHERICHIA COLI FROM A SYNTHETIC GENE" GENE, NL, ELSEVIER BIOMEDICAL PRESS. AMSTERDAM, vol. 136, 22 December 1993 (1993-12-22), pages 323-328, XP000199775 ISSN: 0378-1119 the whole document</p>
A	<p>ANDRE S ET AL: "INCREASED IMMUNE RESPONSE ELICITED BY DNA VACCINATION WITH A SYNTHETIC GP120 SEQUENCE WITH OPTIMIZED CODON USAGE" JOURNAL OF VIROLOGY, US, THE AMERICAN SOCIETY FOR MICROBIOLOGY, vol. 72, no. 2, 1 February 1998 (1998-02-01), pages 1497-1503, XP002073767 ISSN: 0022-538X the whole document</p>